USE OF GAUSSIAN MIXTURE DISTRIBUTION MODELS TO ADDRESS NON-GAUSSIAN ERRORS IN RADAR TARGET TRACKING

A Dissertation
Presented to
The Academic Faculty

by

Benjamin P Davis

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Electrical and Computer Engineering

Georgia Institute of Technology
December 2017

USE OF GAUSSIAN MIXTURE DISTRIBUTION MODELS TO ADDRESS NON-GAUSSIAN ERRORS IN RADAR TARGET TRACKING

Approved by:

Dr. Erik Verriest, Advisor
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Aaron Lanterman
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. William Dale Blair
Georgia Tech Research Institute
Georgia Institute of Technology

Dr. Mary-Ann Weitnauer
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Magnus Egerstedt
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Yao Xie
School of Industrial and Systems Engineering
Georgia Institute of Technology

Date Approved: 8 Nov 2017

To my family,

without whose love and support I would not find myself

in my present happy circumstance.

ACKNOWLEDGEMENTS

I would first of all like to thank my advisor and co-advisor, Dr. Erik Verriest and Dr. Dale Blair. When I first took Dr. Verriest's class in my undergraduate days, I was struck by the depth of mathematical knowledge and rigor he possessed and the passion and clarity with which he shared it. As my department advisor, he has been nothing but supportive in providing feedback and guidance to my work with Dr. Blair. Dr. Blair of GTRI has proved invaluable to my PhD experience. He provided key assistance in directing my attentions to a problem within radar tracking which I felt I had the passion and the ability to address, and he has worked tirelessly to ensure I present my research results in the most professional and effective way possible. I would also like to thank the rest of my present and former GTRI colleagues who provided me personal and professional support along the path of my PhD journey including but not limited to Dr. George Brown, Dr. Norm Speakman, Darrell Wright, Jeff Smart, Dr. John Glass, Dr. Peter Tuuk, Dr. Hunter Hughes, and Dr. Lisa Ehrman. I would like to thank my family, whose patience with me in completing this endeavor has been unbounded. Finally, I would like to thank my friends and roommates who put up with my cryptic discussions on my research and never stopped encouraging me to finish, including David Phillips, Jared Rowland, Josh Wesloski, and Alex Davis. I count myself blessed to have such a supportive personal and professional community.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

The job of a radar tracker is to provide accurate estimates of the kinematic state of a target in order to provide situational awareness, cue other systems, and assist in associating subsequent measurements to the track. The standard algorithm used for many decades by radar trackers is based on the Kalman filter, which requires that the radar measurements be related to the state vector by a linear transformation. However, in most real-world situations this is not the case. Use of the Extended Kalman Filter (EKF), which relaxes this linearity requirement and uses a linearized transformation, can provide useful, sub-optimal state estimates in many cases. This dissertation addresses a well-known situation where the linearization approach fails known as the "contact-lens" problem.

If radar measurements are performed in the polar coordinates of range and angle, and the error orthogonal to the range dimension (i.e., the crossrange) is much greater than the error in range, then the true error region in Cartesian space is no longer well approximated by a Gaussian distribution. This effect is known as the "contact-lens" effect due to the shape of the error distribution in Cartesian space. The true moments of the converted measurement distribution cannot be effectively approximated by linearization of the polar-to-Cartesian transformation. A common approach used to address this issue known as the "debiased transform" is to use higher order formulas to correctly moment match a Gaussian distribution to the measurement. However, this causes problems in two significant ways. First, the Cartesian covariance in the range dimension is inflated compared to the original measurement range covariance, which causes degraded track estimates in range. Furthermore, the gating region that results from the Gaussian covariance ellipse no longer efficiently describes the containment

region of the probability distribution mass. This degrades the measurement-to-track association process, as a significant amount of area with low density is contained within the moment matched Gaussian ellipse, which increases the opportunity for measurement misassociation.

In this dissertation, a method is presented for modeling Cartesian converted measurement distributions which suffer from the contact-lens effect using Maximum Likelihood (ML) Gaussian mixture (GM) parameters. In order to allow an efficient implementation of this process in a GM Kalman filter, a novel normalization of the ML parameters is introduced so that parameters can be efficiently stored in a lookup table for real-time use. Additionally, the measurement update process in the resulting GM filter is modified using a preconditioning process so that the GM measurement PDF is located in close proximity to the support of the state estimate PDF. This preconditioning allows fewer GM components to be used in the model, which significantly reduces the computation costs of the tracking. These techniques are then combined into the Measurement-Adaptive Gaussian Mixture Filter (MAGMF).

The MAGMF approach is first applied to two-dimensional monostatic radar. This is the most straightforward of the problems addressed here, and it has been studied and reported in the literature extensively. Subsequently, two-dimensional bistatic radar measurements are addressed, which provide a greater challenge due to the more complex elliptical geometry. Furthermore, the bistatic radar contact-lens problem has not been as extensively studied in the literature, and no convenient metrics to evaluate the severity of the nonlinearity have so far been developed. Finally, the three-dimensional monostatic radar tracking problem is addressed. This presents a greater difficulty because the ML GM fits to a 3D radar measurement in Cartesian space have widely varying geometries compared to ML fits to a 2D measurement. Novel strategies for storing these ML fits are developed, and these fits are integrated into a 3D MAGMF. The performance of the MAGMF filter is then compared to existing

state-of-the-art filtering methods that attempt to solve the contact-lens problem, including other GM approaches and particle filters. For all three of these cases, the MAGMF is shown to have track accuracy and covariance consistency performance comparable to solutions that use a particle filter that requires significantly more computations.

CHAPTER I

INTRODUCTION

Recursive Bayesian estimation is a widely used technique for state and parameter estimation given noisy input measurements and a stochastic model of system dynamics. In the area of sensing and radar systems, recursive Bayesian estimation is used primarily for tracking targets. Given an estimate of a target's present kinematic state, new measurements are selected to be associated with the track based on their likelihood under the state estimate. Any measurements associated with the track are then used to update the target's state estimate. The classic Kalman filter algorithm [2],[17] provides an optimal solution to the state estimation problem under the conditions of a linear state dynamics process, a linear measurement model, and additive white Gaussian noises in the dynamics and measurements. The Kalman filter involves a simple iteration of prediction and update steps that provide gradual refinement of the state estimate as new measurements arrive. Furthermore, it requires modest computing power.

Unfortunately, the required conditions for the optimality of the Kalman filter do not hold true for most real-world systems and targets. For example, endo-atmospheric ballistic motion where drag forces act upon the target has a non-linear state dynamics model, as the drag forces are proportional to the square of target speed. Maneuvering targets such as a turning aircraft exhibit highly correlated deviations from a straight line trajectory rather than random uncorrelated perturbations characterized by a white noise error model. Furthermore, the natural measurement space of most radar systems is in polar or spherical coordinates of range and angles (or angle sines), whereas the state dynamics of targets are most often expressed in Cartesian

coordinates. Although it is possible to convert state dynamics to the measurement coordinate system to avoid this problem, for complex target models beyond a straight-line dynamics (including maneuvering targets) this becomes increasingly unworkable. Therefore, a non-linear transformation between the state and the measurements is necessitated, resulting in a non-Gaussian distribution for the converted measurement.

In general, these non-linear estimation problems present an intractable challenge to solve optimally as the distributions of the optimal estimate of the state are no longer Gaussian. For general non-Gaussian distributions, closed-form expressions for the application of Bayes' rule no longer exist, and the resulting distributions must be evaluated in terms of probability integrals. Fortunately, the Kalman filter algorithm may be extended to operate sucessfully on a linearized version of the state dynamics and measurement functions. This algorithm, known as the "Extended Kalman Filter" (EKF), provides reasonable sub-optimal estimation performance in the presence of mild non-linearities in these functions. However, in highly non-linear cases, the EKF may produce highly degraded state estimates and may report error covariances that do not represent the true error in the track. As both of these aspects of estimation are critical to the ability of a radar system to associate new measurements into track and to support the reliable fusion of its tracks with other sensors, these shortcomings must be addressed.

1.1 Radar Contact-Lens Problem

One notable case of a non-Gaussian measurement arises in the case of a polar radar measurement performed with high range precision in conjunction with low angular precision. In this case, the standard polar-to-Cartesian measurement transform becomes highly non-linear due to the increased support of the error distribution in the cross-range direction compared to the support in the range direction. Due to the

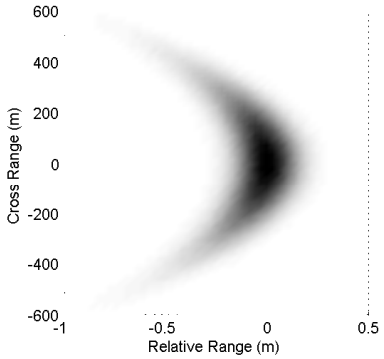

Figure 1: Non-Gaussian "Contact-Lens" Error Distribution for Target at a Range of 200 km with Range Error of 0.1 m and Angle Error of 1 mrad.

non-linearity, transformation of a jointly normal polar error distribution into Cartesian space is poorly approximated by a single Gaussian with mean and covariance estimated by first order Taylor series expansion of the transformation. This problem is known as the radar "contact-lens" problem, and it has been studied in two and three dimensions by numerous authors. An example of the non-Gaussian error region resulting from the polar-to-Cartesian transform is shown in Figure 1. The intensity of the gray color in the figure is proportional to the weight of the PDF at that point in space. It is evident that the distribution shown has a significantly different shape from the Gaussian ellipse commonly used in measurement models.

For mild cases of the radar contact-lens, it is sufficient to approximate the measurement distribution in Cartesian space with a single Gaussian distribution moment matched to the mean and covariance of the true distribution. This is known as the "debiased" or "unbiased" transform [19, 20]. However, in order to maintain consistency, the unbiased transform inflates the distribution covariance in the range dimension resulting in decreased estimation performance in range as well as an oversized covariance ellipse which hinders false alarm rejection and increases the difficulty of measurement-to-track association for closely-spaced objects. In many cases, the estimates of range from the track are worse than the range measurements themselves. Therefore, a key objective of this research is to produce a radar tracking algorithm

that is able to report track estimates with range accuracies that meet or exceed the accuracy of the measurements.

Various other approaches have been applied to mitigate this contact-lens effect, including Uncented Kalman Filter (UKF) [16], adaptive covariance inflation techniques [27], and particle filters [24]. However, these approaches either do not achieve the desired range performance at the start of the track [16, 27], or are extremely expensive computationally [24].

1.2 Technical Approach

Between particle filters and single-Gaussian approximation, a "middle-of-the-road" approach to this problem is to represent the converted measurement distribution with a weighted sum of component Gaussian probability density functions (PDFs), known as a Gaussian mixture (GM). In general, this approach has two challenges. First, an appropriate set of weights, means, and covariances must be selected in real-time for the parameters of the mixture PDF so that it closely approximates the measurement distribution. Second, some technique must be adopted to control the number of components that represent the track state, as otherwise they grow exponentially as new measurements are added. Both of these challenges present significant obstacles to the use of a GM Kalman filter even though the framework for such a filter has existed for decades [26].

The objective of this thesis is to implement a GM filtering framework which mitigates or overcomes these challenges in the contexts of non-linear radar measurements that include 2D monostatic radar in in polar coordinates, 2D bistatic radar in modified elliptical coordinates, and 3D monostatic radar in spherical coordinates. Ideally, this filter will achieve estimation performance comparable to a particle filter with a greatly reduced computation cost and will maintain a conservatively-sized error uncertainty region so that measurement-to-track gating and association may be performed

effectively.

1.3 Summary of Contributions and Organization

This dissertation is organized as follows. Chapter 2 examines the contact-lens problem for 2D monostatic radar, which is the most studied case to date. An information-theoretic interpretation of existing metrics of non-linearity for single-Gaussian models is provided to allow comparison with ML fitting of mixtures. ML fits of GM parameters are performed using the EM algorithm and the results normalized so that they can be stored in an efficient lookup table for use in a real-time GM filter. Additionally, a novel measurement pre-conditioning and inverse filtering approach is presented which allows the number of components in the measurement GM model to be greatly reduced when used in a filter. These techniques are integrated into a GM Kalman filter, the Measurement-Adaptive Gaussian Mixture Filter (MAGMF), which is simulated in comparison to other filtering techniques proposed to handle the contact-lens problem. The results of the simulation confirm that the new filter provides excellent estimation performance with an efficiently-sized gating region at low computational cost compared to a particle filter.

Chapter 3 extends this methodology to 2D bistatic radar, which has a more complex geometry and different measurement non-linearities which are not as well studied in the literature. The key metric of bias significance developed for monostatic radar [19] is extended to apply to bistatic radar as well as the information-theoretic metrics developed in Chapter 2. The lookup tables for ML GM parameters derived for 2D monostatic measurements are also adapted for use in a bistatic radar. The MAGMF is modified to incorporate these bistatic radar measurement models and is simulated against a highly non-linear bistatic tracking scenario. The simulation results show the effectiveness of the filter in bistatic tracking in comparison to single-Gaussian methods and particle filters.

Chapter 4 applies these techniques to 3D monostatic measurements in spherical range-azimuth-elevation (RAE) space. The information-theoretic behavior of ML GM fits to 3D measurements is studied and found to be different than the 2D case. A new measurement metric, the bias significance ratio, is proposed to describe these differences. Measurements with a unity bias significance ratio are found to present the most difficult problem for ML fitting. These measurements are studied extensively in terms of their solution geometries, and a normalization procedure similar to Chapter 2 is developed to allow the parameters to be efficiently stored in a lookup table. Additionally, methods are provided to generate sub-optimal GM solutions for measurements with non-unity bias significance ratio. These models and techniques are integrated into a 3D version of the MAGMF which is simulated in comparison to other proposed approaches from the literature. The new filter is found to have excellent range estimation performance and covariance consistency while having low computational cost. Finally, Chapter 5 provides concluding remarks and recommends future research.

CHAPTER II

GM FILTERING FOR 2D MONOSTATIC RADARS

In this chapter, the contact-lens problem for two-dimensional (2D) monostatic radar systems is considered. This is the simplest formulation of the contact-lens problem possible and the most studied to date. Effective techniques to address the contact-lens problem in the 2D monostatic radar case developed in this chapter will be extended to more complex cases in Chapter 3 and 4.

As described in the introduction, the contact-lens effect causes problems in several significant ways. To begin with, the true moments of the converted measurement distribution cannot be effectively approximated by linearization of the polar-to-Cartesian transformation. Higher order methods are necessary to achieve unbiased and consistent first and second moments. However, even if the moments are calculated correctly, a moment matched Gaussian distribution suffers from an inflated covariance in the range dimension compared to the original range variance of the measurement. This leads to degraded tracker range estimation performance. Furthermore, the gating region defined by the Gaussian covariance ellipse does not efficiently describe the containment region of the distribution mass in Cartesian coordinates. This degrades the measurement-to-track association process, as a significant amount of area with very low probability density is contained within the moment matched Gaussian ellipse, increasing the opportunity for measurement misassociations.

A seminal work on this topic [19] established a measure known as "bias significance" to indicate the severity of the contact-lens problem for a given measurement. If the range error is Gaussian with a mean of $\bar{r}$ and a standard deviation of σ_r, and the angle error is Gaussian with a mean of $\bar{\theta}$ and standard deviation of σ_θ, then bias

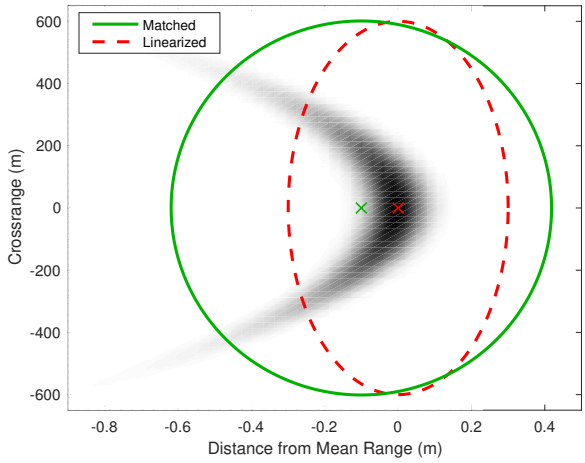

Figure 2: PDF of Converted Measurement with Mean and Covariance Ellipses, $\bar{r} = 200$ km, $\sigma_r = 0.1$ m, $\bar{\theta} = 0$, and $\sigma_\theta = 1$ mrad

significance is given by

$$C_B = \frac{\bar{r}\sigma_\theta^2}{2\sigma_r} \tag{1}$$

In order to illustrate the effect of the nonlinear transformation on the measurement distribution, Figure 2 shows the PDF of a converted measurement with $\bar{r} = 200$ km, $\sigma_r = 0.1$ m, and $\sigma_\theta = 1$ mrad, along with the mean and covariance of two Gaussian approximations using linearized (red) and true (green) moment estimates. The red and green ellipses are 99% containment regions based on the covariance in Cartesian coordinates. The intensity of the gray color in the figure indicates the relative likelihood of the PDF. The bias significance of this measurement is $C_B = 1$. It is seen that the true measurement PDF differs significantly in shape from a Gaussian ellipse. Furthermore, the Cartesian mean computed by linearization (red cross) is biased from the true Cartesian mean in range (green cross), and the linearized covariance (red ellipse) is too small along the range axis compared to the true covariance (green ellipse). The large green covariance ellipse which represents a 99% containment region for the moment matched Gaussian distribution contains a large amount of empty space which will increase measurement-to-track misassociations.

2.0.1 Single-Gaussian Approaches

For bias significance values of $C_B > 0.2$, [19] shows that the standard EKF is covariance inconsistent and yields poor performance. In these cases, a linear Kalman filter in Cartesian measurement space is used to approximate the measurement error using a single Gaussian with mean and covariance matched to the true error distribution. This technique is known as the "debiased" or "unbiased" transform, and it was extended in [20] and [27] to apply to range-azimuth-elevation (RAE) and array face (RUV) measurements. The debiased transform provides near consistent tracking results. However, it degrades the measurement in the range dimension as the covariance is inflated in this dimension to achieve the required moment match. In some cases, the track error in the range dimension may be many times larger than the error reported in the raw measurements.

The work in [27] also proposed an adaptive modification to polar/spherical space EKF known as the Measurement Covariance Adaptive EKF (MCAEKF). This approach consists of inflating the variance of the measurement in range so that a consistency criteria based on the ratio of state and measurement covariance eigenvalues is satisfied. In contrast to the debiased transform, this approach inflates covariance in polar rather than Cartesian space, which sacrifices range performance in order to preserve covariance consistency. However, the adaptive nature of the algorithm allows this inflation to occur only to the extent needed to ensure consistency over the joint measurement-state support region, rather than over the entire measurement support as with the debiased transform.

The Unscented Kalman Filter (UKF) with polar space measurements may also be applied to the contact-lens problem [16, 30]. Due to the fact that the Unscented Transform (UT) correctly matches the distribution moments and the superior accuracy of statistical linearization compared to EKF linearization around the mean, the UKF provides better covariance consistency and estimation performance than the

EKF. However, the UKF also suffers from degraded performance in the range dimension at the beginning of the track, and its covariance is inconsistent for cases with higher values of bias significance.

2.0.2 Alternative Approaches

In general, when applied to measurements with large bias significance, the single-Gaussian approaches examined above suffer from either inconsistent covariance reporting, lack of initial estimation performance in the range dimension, or both. This is due to the fact that a highly non-Gaussian error distribution is being modeled as Gaussian. In order to alleviate this problem, a better representation of the error is found.

One approach used to perform non-Gaussian estimation is particle filtering. In [23, 24] a regularized particle filter (RPF) was applied to the contact-lens problem and achieved consistent covariance with excellent range estimation performance. However, the computational cost of a particle filter is several orders of magnitude higher than a standard Kalman filter as hundreds of thousands of particles may be necessary to provide good results. A particle flow approach suggested by Daum [7] may provide a path to reduce this computational cost, but concrete implementations and results from this approach are currently not in evidence.

Another approach to non-Gaussian filtering is to model the state and/or measurement distribution as a weighted sum of Gaussians, known as a Gaussian mixture (GM). This approach was first suggested by Alspach and Sorenson [1] in the 1970s, but it has traditionally seen limited use due to two major difficulties. First, mixture approximations for a non-Gaussian distribution must be generated by an appropriate method in order to be used in the filter in real-time. The optimal solution to this problem from a Maximum-Likelihood (ML) perspective would be to use Expectation Maximization (EM) or a similar algorithm to find a fit numerically, but this

is intractable for real-time use in a filter due to the slow convergence of the algorithm. Second, when the measurement and state are both represented by a Gaussian mixture, the update step of the Kalman filter results in a mixture that increases in size exponentially with the number of measurements. Therefore, some strategy must be implemented to control this growth, and this has been found to be challenging. However, as computational power of systems has increased over time, the number of applications of GM filters has seen an increase. Some recent application areas include space tracking [15] and robotics [14].

In [28], the Consistency-based Gaussian Mixture Filter (CbGMF) was proposed to address the contact-lens problem. This approach uses an ad-hoc technique to generate a GM representation of the state PDF. It conditions the state prior along various intervals of the cross-range position axis. The number of components in the split is chosen so that a consistency metric of curvature similar to a limit on the bias significance is satisfied for each component. This allows the components to be updated using a standard EKF. As the measurement is represented by a single Gaussian, exponential growth of the representation is not a problem. However, additional splitting of the state may be required due to uncertainty introduced during propagation, and that increases the number of Gaussians required for the state representation. To control the size of the state mixture, the CbGMF periodically prunes mixture components with low likelihood and performs a re-splitting procedure which combines all components back to a single Gaussian as an initial step. Unfortunately, this combining/re-splitting approach degrades the local structure of the state density and reduces performance. Furthermore, when tracking with measurements having a high bias significance, a large number of initial state components may be necessary to achieve good performance.

In [32] and [33], an approach called Gaussian Mixture Measurement-Integrated Track Splitting (GMM-ITS) was proposed which uses a Gaussian mixture to split

the measurement distribution rather than the state. An ad-hoc approach similar to [28] was used to achieve this split. As a result of updating the state with a measurement mixture, the number of components in the filter grows exponentially with the number of measurements. To control the number of components, all posterior state components that result from interaction with a particular measurement component are reduced to a single Gaussian. Additionally, components with low weight are dropped over time. This GM filter achieves good range estimation performance with reasonable consistency. However the number of components required to represent a measurement with high bias-significance is relatively high, which increases the computational cost of the filter. Additionally, during the research for this chapter it was found that the GM fitting method used in GMM-ITS has a fixed lower bound on the Kullback-Leibler (KL) divergence between the GM model and the true distribution, which increases the size of the gating region and degrades range estimation performance.

2.0.3 Proposed Approach

Rather than use ad-hoc methods to generate GM models of a radar measurement, ML is used to optimally determine the GM parameters of weight, mean, and covariance for a given measurement. These ML parameters will be generated using the EM algorithm. However, this approach is not practical for use in an real-time filter, as the ML process is computationally expensive. Therefore, the ML parameters resulting from EM are stored in a lookup table to allow real-time use in a filter. Additionally, by monitoring the KL divergence achieved by the ML fits, the number of GM components needed to achieve reasonable performance can be adaptively chosen. As a result, fewer components are used to model the state and measurement than in the CbGMF and GMM-ITS, resulting in reduced computational cost and equivalent or superior estimation performance.

Initial application of this approach in a Kalman filter revealed that it is fundamentally limited in cross-range performance, since as the track state converges many of the Gaussian mixture components used to model the measurement are no longer located in positions which are useful to the update. This is due to the fact that an ML fit is optimized over the entire support of the measurement distribution, but in the case of a converged track, the support region of interest is much smaller. The result of this is that many of the components in the posterior distribution have very low weight and only a few contribute meaningfully to the likelihood. If this difficulty is not remedied, the filter needs to use a greater number of components to model the measurement in order to ensure sufficient modeling accuracy of the posterior distribution. Therefore, a measurement pre-conditioning technique is used to reduce the mismatch between the state and measurement PDF before generating a mixture model. This conditioning factor is later removed by an inverse filtering process to give an estimate of the posterior state distribution.

2.0.4 Organization

This chapter is organized as follows. First, key results regarding the KL divergence performance of existing single-Gaussian measurement modeling methods are presented. Second, the fitting procedure used to generate an efficiently-sized lookup table of ML results is described. Third, the difficulties presented by track convergence are discussed and a solution based on strategic conditioning of the measurement model is presented. Fourth, this conditioning solution is applied to a GM Kalman filter, resulting in the new Measurement-Adaptive Gaussian Mixture Filter (MAGMF) algorithm. Finally, performance of the MAGMF is compared to recently proposed solutions to the contact-lens problem. These include the MCAEKF, CbGMF, GMM-ITS, and RPF.

2.1 Overview of the 2D Contact-Lens Problem

A two-dimensional monostatic radar system makes measurements in polar range and angle space. A common model for the errors in these measurements is independent Gaussian measurement noises in both range and angle. Let $w_r \sim N(0, \sigma_r^2)$ and $w_\theta \sim N(0, \sigma_\theta^2)$, where $N(\mu, \sigma^2)$ indicates a normal (Gaussian) distribution with mean μ and variance σ^2. Then the 2D measurements with errors are given by

$$r = \bar{r} + w_r \tag{2}$$

$$\theta = \bar{\theta} + w_\theta \tag{3}$$

where $\bar{r}$ and $\bar{\theta}$ represent the true (mean) values of the measurements.

These measurements can be converted to Cartesian space by the following transformation.

$$x = r\cos(\theta) \tag{4}$$

$$y = r\sin(\theta) \tag{5}$$

It can be assumed that $\bar{\theta} = 0$ without loss of generality, since a rotation matrix can be applied to the Cartesian axes to remove any non-zero mean angle. Therefore, any results determined for $\bar{\theta} = 0$ can be rotated to an arbitrary $\bar{\theta}$.

The moments of this transformation are approximated with a first order Taylor expansion as

$$E\{x\} \approx \bar{r} \tag{6}$$

$$E\{y\} \approx 0 \tag{7}$$

$$Cov\left\{\begin{bmatrix} x \\ y \end{bmatrix}\right\} \approx \begin{bmatrix} \sigma_r^2 & 0 \\ 0 & \bar{r}^2\sigma_\theta^2 \end{bmatrix} \tag{8}$$

In many cases, this linearized approach is adequate to model the Cartesian measurement distribution. However, if the range error becomes very small in comparison to the cross-range error, this approximation of the distribution moments becomes inaccurate as illustrated in Figure 2.

2.1.1 Bias Significance

The bias significance measure previously described in (1) assists in quantifying the severity of the contact-lens problem. The origin of this measure can be understood by examining the expansion of the polar-to-Cartesian conversion (4) and (5). From [19], $E\{\cos(\theta)\} = e^{-\frac{\sigma_\theta^2}{2}}$ and $E\{\sin(\theta)\} = 0$. Thus,

$$E\{x\} = \bar{r}e^{-\frac{\sigma_\theta^2}{2}} \tag{9}$$

$$E\{y\} = 0 \tag{10}$$

In the linearized approximation, $E\{x\} \approx \bar{r}$. Therefore, the "bias" introduced by the linearization of the transformation is $\bar{r}\left(1 - e^{-\frac{\sigma_\theta^2}{2}}\right) = \bar{r}\frac{\sigma_\theta^2}{2} + O(\sigma_\theta^4)$. This bias becomes significant if it is comparable in magnitude to the range standard deviation σ_r. Thus, the stated bias significance equation in [19] represents the bias between the true and linearized mean normalized by the range standard deviation. The bias significance also affects the covariance calculation. Given

$$E\left\{\cos^2(\theta)\right\} = \frac{1}{2}\left(1 + e^{-2\sigma_\theta^2}\right) \tag{11}$$

$$E\left\{\sin^2(\theta)\right\} = \frac{1}{2}\left(1 - e^{-2\sigma_\theta^2}\right) \tag{12}$$

$$\begin{aligned} Var\{x\} &= E\left\{r^2\right\} E\left\{\cos^2(\theta)\right\} - E\{r\}^2 E\{\cos(\theta)\}^2 \\ &= \frac{1}{2}\left(\bar{r}^2 + \sigma_r^2\right)\left(1 + e^{-2\sigma_\theta^2}\right) - \bar{r}^2 e^{-\sigma_\theta^2} \\ &= \sigma_r^2 + \frac{1}{2}\bar{r}^2\sigma_\theta^4 + O\left(\sigma_r^2\sigma_\theta^2, \bar{r}^2\sigma_\theta^6\right) \\ &= \sigma_r^2\left(1 + 2C_B^2\right) + O\left(\sigma_r^2\sigma_\theta^2, \bar{r}^2\sigma_\theta^6\right) \end{aligned} \tag{13}$$

$$\begin{aligned} Var\{y\} &= E\left\{r^2\right\} E\left\{\sin^2(\theta)\right\} \\ &= \frac{1}{2}\left(\bar{r}^2 + \sigma_r^2\right)\left(1 - e^{-2\sigma_\theta^2}\right) \\ &= \bar{r}^2\sigma_\theta^2 - 2\bar{r}^2\sigma_\theta^4 + O\left(\sigma_r^2\sigma_\theta^2, \bar{r}^2\sigma_\theta^6\right) \\ &= \bar{r}^2\sigma_\theta^2\left(1 - \sigma_\theta^2\right) + O\left(\sigma_r^2\sigma_\theta^2, \bar{r}^2\sigma_\theta^6\right) \end{aligned} \tag{14}$$

The variances from first-order linearization are $Var\{x\} \approx \sigma_r^2$ and $Var\{y\} \approx \bar{r}^2\sigma_\theta^2$. Therefore, the true X (range) variance is inflated from the linearized version by a factor of approximately $1 + 2C_B^2$. The Y (crossrange) variance is adjusted by a factor of $(1 - \sigma_\theta^2)$, which is negligible if $\sigma_\theta \ll 1$ (true in most all cases).

2.1.2 Alternative Curvature Measures

A number of curvature measures other than bias significance are in the literature. In [29], a "curvature" parameter is used to assess the consistency of the polar measurement transformation based on the maximum distance of the Gaussian distribution to the iso-range curve of the true measurement. The distance $D_C^r(x)$ of a given point x in the true uncertainty region from the estimated Gaussian ellipse is defined as

$$D_C^r(x) = \frac{r_e^2}{r_p(x)} - r_e \tag{15}$$

where $r_e = |\bar{x}|$, $r_p(x) = \bar{x}\dot{x}|x|^{-1}$, and $\bar{x}$ is the mean of the ellipse. The curvature parameter is then given by

$$c_a^r = \frac{max_{x\in\Phi_\alpha} D_C^r(x)}{\sigma_r} \tag{16}$$

where Φ_α is the $\alpha\%$ confidence region of the ellipse. Note that $r_p {r_e}^{-1} = \cos\theta$, where θ is the angle between the mean and the chosen point on the sphere (the angle of the mean can be chosen as zero). Then

$$D_C^r(x) = r_e\left(\frac{1}{\cos\theta} - 1\right) \tag{17}$$

The maximum of this value occurs when the cosine term is minimized or when θ is maximized within Φ_α. If ϕ_α is the Chi-squared score which leads to a CDF value of α for a Chi-squared distribution of two degrees of freedom, then

$$c_a^r = \frac{r_e}{\sigma_r}\left(\frac{1}{\cos\sigma_\theta\sqrt{\phi_\alpha}} - 1\right) \tag{18}$$

Using the expansion for the secant $\frac{1}{\cos x} = 1 + \frac{1}{2}x^2 + o(x^3)$ and using $r_e = \bar{r}$ as an equivalent expression of the mean gives

$$c_a^r = \phi_\alpha\frac{\bar{r}\sigma_\theta^2}{2\sigma_r} \tag{19}$$

which is equivalent to a scaled version of the bias significance in [19]. Tian uses a limit of $c_a^r \leq \frac{1}{4}$ along with a 2σ containment region for α in [29], which leads to $\phi_\alpha = 6$ for the Chi-squared score. This is equivalent to limiting the bias significance to $C_B \leq \frac{1}{24}$, which is well within the limit of 0.2 prescribed by [19]. Therefore, the CbGMF filter's "consistency rule" may be re-formulated in terms of a bias significance limit.

2.2 Information-Theoretic Metrics for Estimation Performance

Kullback-Leibler (KL) divergence is a quantity used in information theory to express the loss of performance in a sub-optimal coding scheme [18]. In general, it is a quantity that expresses the difference in information between two distributions. Suppose a message with probability distribution $p(x)$ is transmitted using a code designed for the distribution $q(x)$. Then the number of extra units needed to transmit the message is given by

$$D_{KL}(P||Q) = \int_{-\infty}^{\infty} p(x) \ln \frac{p(x)}{q(x)} \, dx \tag{20}$$

which is equivalent to the expectation of $\ln p(x)q(x)^{-1}$ taken over $p(x)$. The units of divergence are "bits" if the logarithm is taken to base 2 and "nats" if the logarithm is taken to base e. Given the exponential form of the normal distribution PDF, it is convenient to use the natural logarithm.

Additionally, the the differential cross-entropy of p and q is given by

$$H\left(p,q\right) = -\int_X p\left(x\right) \ln q\left(x\right) \, dx \tag{21}$$

and differential entropy of p is given by

$$H\left(p\right) = -\int_X p\left(x\right) \ln p\left(x\right) \, dx \tag{22}$$

By the properties of the logarithm, the KL divergence may be written as the difference in cross-entropy and entropy.

$$D_{KL}\left(P||Q\right) = H\left(p,q\right) - H\left(p\right) \tag{23}$$

2.2.1 Entropy of a Gaussian Random Variable

The entropy of a multi-dimensional Gaussian (in nats) has a convenient closed form which is seen by examining the PDF function [22]

$$N\left(x;\mu,\Sigma\right)=\frac{1}{\sqrt{\left(2\pi\right)^{k}\left|\Sigma\right|}}\exp\left(-\frac{1}{2}(x-\mu)^{T}\Sigma^{-1}\left(x-\mu\right)\right) \tag{24}$$

with

$$\ln\left(N\left(x;\mu,\Sigma\right)\right)=-\frac{k}{2}\ln\left(2\pi\right)-\frac{1}{2}\ln\left|\Sigma\right|\ -\frac{1}{2}(x-\mu)^{T}\Sigma^{-1}\left(x-\mu\right) \tag{25}$$

Since $(x-\mu)^{T}\Sigma^{-1}\left(x-\mu\right)$ is a Chi-squared random variable with k degrees of freedom, $E\left\{(x-\mu)^{T}\Sigma^{-1}\left(x-\mu\right)\right\}=k$ which leads to

$$H\left(p\right)=\frac{k}{2}\left(1+\ln\left(2\pi\right)\right)+\frac{1}{2}\ln\left|\Sigma\right| \tag{26}$$

2.2.2 Cross-Entropy of a Transformed Gaussian

Suppose $p_x(x)\sim N\left(x;\mu,\Sigma\right)$ is transformed by a function to yield $y=f\left(x\right)$. By transformation of variables [22],

$$p_{y}\left(y\right)=\frac{p_{x}\left(f^{-1}\left(y\right)\right)}{\left|\frac{\partial f}{\partial x}\right|} \tag{27}$$

Now, let $p_y(y)$ be modeled by a distribution $q_y(y)$ (for example a Gaussian). The equivalent distribution in x is

$$q_{x}\left(x\right)=\left|\frac{\partial f}{\partial x}\right|q_{y}\left(f\left(x\right)\right) \tag{28}$$

Then the cross-entropy in x is

$$\begin{aligned}H\left(p,q\right)&=-\int_{X}p_{x}\left(x\right)\ln q_{x}\left(x\right)dx\\&=-\int_{X}p_{x}\left(x\right)\ln\left(\left|\frac{\partial f}{\partial x}\right|q_{y}\left(f\left(x\right)\right)\right)dx\\&=-E_{p}\left\{\ln q_{y}\left(f\left(x\right)\right)\right\}-E_{p}\left\{\ln\left|\frac{\partial f}{\partial x}\right|\right\}\end{aligned} \tag{29}$$

This convenient equation allows the cross-entropy to be calculated in the original integration space (x) even if the model distribution is represented in alternative coordinates.

2.2.3 KL Divergence of 2D Monostatic Contact-Lens

Here, the KL divergence is calculated for two single-Gaussian models of the converted measurement described by (4) and (5). It is seen that for moderate angular error the KL divergence is approximated as a direct function of the bias significance and uses no other independent parameters of the the original distribution. In order to calculate the cross-entropy as described above, the Jacobian of the transformation is calculated according to

$$\left|\frac{\partial f}{\partial x}\right| = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r \tag{30}$$

Using (26), the entropy of the original Gaussian measurement is

$$H\left(p\right) = 1 + \ln\left(2\pi\right) + \ln\left(\sigma_r \sigma_\theta\right) \tag{31}$$

2.2.3.1 KL Divergence of First-order Converted Measurements Model

If (6) through (8) are used to calculate the moments of the transformed distribution and

$$q_{\boldsymbol{y}}\left(\boldsymbol{y}\right) = N\left(\boldsymbol{y} = \begin{bmatrix} x \\ y \end{bmatrix}; \mu_q, \Sigma_q\right) \tag{32}$$

If

$$\mu_q = \begin{bmatrix} \bar{r} \\ 0 \end{bmatrix} \tag{33}$$

$$\Sigma_q = \begin{bmatrix} \sigma_r^2 & 0 \\ 0 & \bar{r}^2\sigma_\theta^2 \end{bmatrix} \tag{34}$$

are used to model the distribution, then the cross-entropy $H(p,q)$ is given by

$$H(p,q) = \ln 2\pi + \ln\left(\sigma_\theta \sigma_r\right) - E\left\{\ln\frac{r}{\bar{r}}\right\} + E\left\{\frac{1}{2}\left(\boldsymbol{y}-\mu_q\right)\Sigma_q^{-1}\left(\boldsymbol{y}-\mu_q\right)^T\right\} \tag{35}$$

For $\sigma_r \ll r$ (which is true for almost all radars), $E\left\{\ln\frac{r}{\bar{r}}\right\} \approx 0$. It now remains to calculate the Mahalanobis distance. Let $M = \left(\boldsymbol{y}-\mu_q\right)\Sigma_q^{-1}\left(\boldsymbol{y}-\mu_q\right)^T$ for brevity.

Then

$$\begin{aligned} E\{M\} &= E\left\{\frac{\left(r\cos\theta-\bar{r}\right)^2}{\sigma_r^2}\right\}+E\left\{\frac{\left(r\sin\theta\right)^2}{\bar{r}^2\sigma_\theta^2}\right\} \\ &= \frac{3}{4}\frac{\bar{r}^2\sigma_\theta^4}{\sigma_r^2}+2\left(1-\sigma_\theta^2\right)+O\left(\bar{r}^2\sigma_\theta^6,\sigma_\theta^4,\frac{\sigma_r^2}{\bar{r}^2}\right) \\ &= 3C_B^2+2\left(1-\sigma_\theta^2\right)+O\left(\bar{r}^2\sigma_\theta^6,\sigma_\theta^4,\frac{\sigma_r^2}{\bar{r}^2}\right) \end{aligned}$$

Since $\sigma_\theta^2 << 1$, the final result is

$$D_{KL}\left(P||Q\right)=\frac{3}{2}C_B^2+O\left(\bar{r}^2\sigma_\theta^6,\sigma_\theta^4,\frac{\sigma_r^2}{\bar{r}^2}\right) \tag{36}$$

2.2.3.2 KL Divergence of Moment-matched Converted Measurements Model

If the debiasing procedure described in [19] is applied, the model distribution for the measurement is given by

$$\begin{aligned} q_{\boldsymbol{y}}\left(\boldsymbol{y}\right) &= N\left(\boldsymbol{y};\mu_q,\Sigma_q\right) \\ \mu_q &= \begin{bmatrix} \bar{r}e^{-\frac{\sigma_\theta^2}{2}} \\ 0 \end{bmatrix} \end{aligned} \tag{37}$$

$$\Sigma_q = \begin{bmatrix} \sigma_r^2\left(1+2C_B^2\right) & 0 \\ 0 & \bar{r}^2\sigma_\theta^2 \end{bmatrix}+O\left(\frac{\sigma_r^2}{\bar{r}^2},\sigma_\theta^2,C_B^2\sigma_\theta^2\right) \tag{38}$$

The cross-entropy is

$$\begin{aligned} H(p,q) =& \ln 2\pi+\frac{1}{2}\ln\left(\sigma_\theta^2\sigma_r^2\left[1+2C_B^2\right]\right)+ \\ & E\left\{\frac{1}{2}\left(\boldsymbol{y}-\mu_q\right)\Sigma_q^{-1}\left(\boldsymbol{y}-\mu_q\right)^T\right\}-E\left\{\ln\frac{r}{\bar{r}}\right\}+ \\ & O\left(\ln\frac{\sigma_r^2}{\bar{r}^2},\ln\sigma_\theta^2,\ln C_B^2\sigma_\theta^2\right) \end{aligned} \tag{39}$$

Since the moments of the distribution are matched,

$$E\left\{\left(\boldsymbol{y}-\mu_q\right)\Sigma_q^{-1}\left(\boldsymbol{y}-\mu_q\right)^T\right\}=2$$

Therefore, the KL divergence of the moment-matched Gaussian distribution is

$$D_{KL}\left(P||Q\right)=\frac{1}{2}\ln\left(1+2C_B^2\right)+O\left(\ln\frac{\sigma_r^2}{\bar{r}^2},\ln\sigma_\theta^2\right) \tag{40}$$

By moment-matching the Gaussian distribution using the debiased transform, the dependence of KL divergence on bias significance is moved inside a logarithm, resulting in a significantly smaller divergence.

2.2.4 Significance of KL Divergence Results

As a result of the above derivation, it is now possible to relate the ad-hoc measures of performance described by the bias significance of [19] and the curvature of [29] to the widely-used information-theory measure of KL divergence. The guidance suggested in [19] that the EKF should not be used if $C_B > 0.2$ may be converted to a KL divergence limit using (36) and (40). A bias significance of $C_B = 0.2$ corresponds to a KL divergence of 0.06 when the linearized moment estimate is used and a KL divergence of 0.038 when the moment match is used. These single-Gaussian KL divergence levels provide a useful heuristic for choosing the limits of KL divergence allowed when a GM is used to model the measurement in order to retain good performance.

A geometric interpretation of the KL divergence results is possible. Consider that the non-zero terms contributing to the KL divergence in the moment matched case are

$$D_{KL}\left(P||Q\right) = \frac{1}{2}\ln\frac{|\Sigma_q|}{|\Sigma_p|} - E\left\{\ln\frac{r}{\bar{r}}\right\} \tag{41}$$

where the last term is neglected for small σ_r.

The area within the 1-sigma (Malahanobis distance = 1) boundary of a covariance ellipse is π times the product of its eigenvalues. Therefore, the square root of the determinant of a covariance matrix is proportional to the area inside the ellipse it describes. The KL divergence expression in the moment-matched case is essentially a statement about the relative area in the covariance ellipses of the model and true distributions. If the area in the model covariance is much larger than the area in the true covariance (where the polar and Cartesian areas are scaled appropriately by the Jacobian of the transform), this "wasted" area can be linked directly to a

loss of information expressed by the KL divergence (which leads to degraded range estimation performance). The above equation states that the scale factor for the area is

$$\sqrt{\frac{|\Sigma_q|}{|\Sigma_p|}} \approx \sqrt{1 + C_B^2} \tag{42}$$

Therefore, if $C_B = 1$, the area in the Cartesian measurement covariance ellipse is $\sqrt{2}$ times the area in the polar measurement (after accounting for the integration factor). This represents an increased opportunity for spurious measurements to be associated with the track.

2.3 Gaussian Mixture Models for Contact-Lens

The "debiased transform" approach to the contact-lens problem is effective for approximating moderately non-Gaussian measurements by a single Gaussian. However, (40) predicts a significant loss of information even for a moment-matched Gaussian model if the bias significance is high. To preserve this information, an alternative model is needed.

In [1], it is noted that a non-Gaussian distribution can be approximated to arbitrary accuracy by a GM of sufficient size. Therefore, this approach will be adopted to decrease the KL divergence of the Cartesian model distribution from the measurement distribution.

A GM is a weighted sum of N_G individual Gaussian *component* densities with PDF given by

$$q(x) = \sum_{k=1}^{N_G} \omega_k N\left(x; \mu_k, \Sigma_k\right) \tag{43}$$

where $\omega_k \geq 0$ is the *component probability weight*, and $N\left(x; \mu, \Sigma\right)$ indicates the PDF of a Normal distribution with mean μ and covariance Σ. Let $\Theta = \{\omega_k, \mu_k, \Sigma_k\}_{k=1}^{N_G}$ denote the parameters of the GM.

For a two dimensional domain, Θ represents a set of $6N_G$ parameters. The radar system supplies only mean and variance of range and angle. Furthermore, the weights

of the mixture ω_k must sum to 1. This results in 6 input constraints on $6N_G$ output parameters, which is an underdetermined problem for $N_G > 1$. The key question is how to provide additional constraints to choose the rest of the parameters.

If an ML approach is used, Θ is chosen to minimize the KL divergence of the model from the true distribution. This requires optimizing

$$\min_{\Theta} \int p(x) \ln \left(\sum_{k=1}^{N_G} \omega_k N\left(x; \mu_k, \Sigma_k\right) \right) dx \tag{44}$$

which contains the logarithm of a sum. In general, this optimization does not have a closed-form solution. Therefore, the ML process will be approached using numerical methods.

2.3.1 EM Fitting of Gaussian Mixture

A popular method for finding ML fits of distribution parameters for discrete data is the EM algorithm. Suppose $X = \{x_i\}_{i=1}^{N}$ are the sample data, and $q(X; \Theta, \Psi)$ is the model PDF function. It is desired to optimize the parameters of interest Θ given the existence of additional unknown (nuisance) parameters Ψ [3, 13].

Given an initial guess Θ for the parameters of interest, the following two steps are iterated.

1. Expectation: Find the expected value of the nuisance parameters Ψ given the data and the current estimate Θ.

2. Maximization: Updated the parameter estimate to Θ^{new} by finding ML estimates given the data and current expected nuisance parameters Ψ.

The algorithm is iterated until the desired level of convergence of the parameters is achieved.

For a GM model, the nuisance parameters are the assignments of samples to a given mixture component. These are expressed as $z_k(x_i)$, where $z_k(x_i) = 1$ if component k is "responsible" for a given sample x_i. The expectation step is to

find the expected probability of a sample's assignment to each component given the current parameters Θ. These are known as the "responsibility" values.

$$r_{i,k} = E\left\{z_k(x_i)|\Theta\right\} = \frac{N\left(x_i;\mu_k,\Sigma_k\right)}{\sum_{j=1}^{N_G} N\left(x_i;\mu_j,\Sigma_j\right)} \tag{45}$$

The maximization step then finds ML estimates of the parameters given the current responsibility values. This can be done by performing ML on each individual Gaussian component, where the input samples to the ML equation are weighted according to the "responsibility" of the given component for each sample.

$$\omega_k^{new} = \frac{1}{N}\sum_{i=1}^{N} r_{i,k} \tag{46}$$

$$\mu_k^{new} = \frac{\sum_{i=1}^{N} r_{i,k} x_i}{\sum_{i=1}^{N} r_{i,k}} \tag{47}$$

$$\Sigma_k^{new} = \frac{\sum_{i=1}^{N} r_{i,k}(x_i - \mu_k^{new})^T(x_i - \mu_k^{new})}{\sum_{i=1}^{N} r_{i,k}} \tag{48}$$

2.3.2 Use of EM Algorithm to Generate GM Models of Contact-Lens

The EM algorithm can be applied to a sampled version of the true measurement distribution to generate Gaussian mixture parameters. However, in order to achieve reasonable accuracy, a large number of samples must be used, resulting in a large computational cost. Therefore, EM cannot be reliably used as an online strategy for generating Gaussian mixtures in a real-world system. Some method of storing the results of EM for quick recall is necessary.

Monte-Carlo runs of EM against sampled radar measurements are made where the input parameters N_G, $\bar{r}$, $\bar{\theta}$, σ_r, and σ_θ are varied and the output parameters Θ recorded in a lookup table. This allows Θ for a given set of inputs to be looked up on demand and used in an online filter. However, the number of runs required to accurately estimate the values of Θ over the space of useful combinations of five input parameters is not tractable. Furthermore, this results in a very large lookup table.

Therefore, it is necessary to reduce the dimension of the input parameter space to a size for which a table of results can be practically stored.

First, the mean angle $\bar{\theta}$ is discounted as an input parameter as it is possible to transform a distribution with a zero-mean angle to an arbitrary mean angle by use of a rotation matrix which preserves angles and areas of the geometry. Additionally, it is expected that the overall range "scale" of the problem will likewise affect the resulting parameter set Θ only by multiplicative factors as long as $\frac{\bar{r}}{\sigma_r}$ is held constant. It is further shown that there exists a transformation of the parameters Θ which is constant for a constant $C_B = \frac{\bar{r}\sigma_\theta^2}{2\sigma_r}$. This property reduces the space of necessary Monte-Carlo runs of EM to the two dimensions of bias significance C_B and number of components N_G, which make a reasonable set of results to be stored in a lookup table.

2.3.3 KL Divergence Performance of EM Fits

The EM algorithm performs ML estimation of the parameters Θ given data samples X. The ML acts to minimize the KL divergence of the model $q(X;\Theta,\Psi)$ from the samples or equivalently, the discrete cross-entropy $H(q,X)$. In the general case, the cross-entropy of a Gaussian mixture from the input set of samples is given by

$$H(q,X) = -\sum_{i=1}^{N} \ln\left(\sum_{k=1}^{N_G} \omega_k N\left(x_i;\mu_k,\Sigma_k\right)\right) \tag{49}$$

For the contact-lens problem, a set of samples $X' = \{(r_i,\theta_i)\}_{i=1}^{N}$ is first generated in polar space by sampling the original Gaussian measurement distribution. These samples are then transformed into Cartesian samples X by application of the polar-to-Cartesian transform in (4) and (5). A Gaussian mixture model $q(X;\Theta,\Psi)$ is then generated by EM. The original entropy Gaussian $H(p)$ used to generate X' is given by (31). The cross entropy $H(q,X)$ can be calculated numerically using (49). In order to find the KL divergence, the cross-entropy and entropy must be represented in the same integration space, so $H(q,X)$ must be converted to $H(q,X')$ by application of

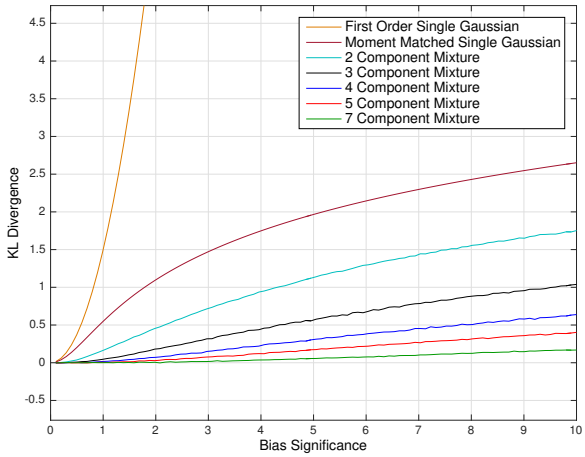

Figure 3: KL Divergence of Gaussian Mixture Fits Compared to Single Gaussian Models

the transform Jacobian term described in (30). Therefore, the KL Divergence of an ML Gaussian mixture fit from the true sampled distribution is given by

$$D_{KL}\left(p||q\right) = -\sum_{i=1}^{N} \ln\left(\sum_{k=1}^{N_G} \omega_k N\left(x_i; \mu_k, \Sigma_k\right)\right) - \sum_{i=1}^{N} \ln r_i - H(p) \tag{50}$$

where r_i are the range samples generated in X'.

Section 2.2 shows that the KL divergence of single-Gaussian models of a polar radar measurement from the true distribution are constant with constant bias significance (C_B). By investigating the KL divergence of ML GM fits using (50), this was found to also be the case for GM models of the measurement. Given an ML estimate of the parameters of GM model of given size N_G for a radar measurement with a given C_B, then the KL divergence and component weight parameters ω_k are the same (within sampling error) independent of changes in the individual input parameters $\bar{r}$, $\bar{\theta}$, σ_r, and σ_θ. This is very significant, as it suggests that the parameter estimates are fundamentally similar, and that it should be possible to transform the component means μ_k and covariances Σ_k such that these parameters are also the same for a given C_B and N_G.

Figure 3 shows the KL divergence estimates of the GMs with ML-estimated parameters for varying size N_G and bias significance C_B, along with the KL divergence

Table 1: Maximum Bias Significance of Models for $KL \leq 0.06$

Model	Max C_B
First Order SG	0.2
Moment Matched SG	0.25
2 Component GM	0.6
3 Component GM	1.1
4 Component GM	1.8
5 Component GM	2.7
6 Component GM	3.7
7 Component GM	5.0
8 Component GM	6.5
9 Component GM	8.2
10 Component GM	10

of the single-Gaussian models studied in Section 2.2. The results show that the KL divergence of the model distribution from the true measurement can be greatly decreased by the use of a GM of appropriate size compared to the performance achieved by single-Gaussian approximations. By using the result of Section 2.2 that the KL divergence of a linearized single-Gaussian model with $C_B = 0.2$ is 0.06, the KL divergence curves can be used to suggest the appropriate N_G for a given C_B. Table 1 summarizes the maximum C_B that should be used for each model order to maintain a KL divergence of at most 0.06. ($N_G = 6$ was not included in the plot for clarity, but it is included in the table.) The linearized single-Gaussian model very quickly reaches a high KL divergence and is unsuitable for use with any appreciable bias significance. If the debiased transform (or alternatively Unscented transform) is used to moment match a single Gaussian model, the KL divergence is improved in asymptotic behavior compared to the linearized model, but the KL limit of 0.06 is still reached by $C_B = 0.25$, which is only a slight improvement over the linearized case of $C_B = 0.2$. Beyond this limit, significant range variance inflation will occur. In contrast, GM

models can be chosen which handle cases up to $C_B = 10$ with $N_G = 10$ or fewer components while maintaining a KL divergence of 0.06 or less.

2.3.4 Transformation of ML Parameters

As suggested in the previous section, the following set of transformations of the ML component means and covariances result in a set of parameters that is also constant for a given N_G and C_B. In [8] this process was performed in Cartesian space, but later research shows that it is most convenient to convert the components back to the original measurement space as the first step in the transformation.

Let C2P denote the Cartesian to polar transformation function defined as

$$\begin{bmatrix} r \\ \theta \end{bmatrix} = \text{C2P}\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} \sqrt{x^2+y^2} \\ \tan^{-1}\frac{x}{y} \end{bmatrix} \tag{51}$$

so that

$$\begin{bmatrix} r_k \\ \theta_k \end{bmatrix} = \text{C2P}(\mu_k) \tag{52}$$

Then, shift and normalize these polar coordinates as follows

$$\Delta\hat{r}_k = \frac{r_k - \bar{r}}{C_B\sigma_r} \tag{53}$$

$$\Delta\hat{\theta}_k = \frac{\theta_k - \bar{\theta}}{\sigma_\theta} \tag{54}$$

Next, convert the covariances to the measurement space using linearization around each mean

$$\Sigma_{P,k} = J_{\text{C2P}}(\mu_k)\Sigma_k J_{\text{C2P}}^T(\mu_k) \tag{55}$$

where $J_{\text{C2P}}(\mu_k)$ is the Jacobian of C2P evaluated at μ_k.

Finally, scale the covariance matrix according to the original distribution errors by

$$\hat{\Sigma}_{P,k} = S^{-1}\Sigma_{P,k}S^{-T} \tag{56}$$

where

$$S = \begin{bmatrix} \sigma_r\sqrt{1+2C_B^2} & 0 \\ 0 & \sigma_\theta \end{bmatrix} \tag{57}$$

and decompose this matrix into

$$\hat{\sigma}_{r,k} = \sqrt{[\hat{\Sigma}_{P,k}]_{(1,1)}} \tag{58}$$

$$\hat{\sigma}_{\theta,k} = \sqrt{[\hat{\Sigma}_{P,k}]_{(2,2)}} \tag{59}$$

$$\rho_{r\theta,k} = \frac{[\hat{\Sigma}_{P,k}]_{(1,2)}}{\hat{\sigma}_{r,k}\hat{\sigma}_{\theta,k}} \tag{60}$$

The lookup table point for a given bias significance C_B and number of Gaussian components N_G is constructed by storing the vector of parameters ω_k, $\Delta\hat{r}_k$, $\Delta\hat{\theta}_k$, $\hat{\sigma}_{r,k}$, $\hat{\sigma}_{\theta,k}$, and $\rho_{r\theta,k}$, along with the KL divergence of the fit evaluated using (50). Table 2 shows the lookup table parameters calculated for $N_G = 2$. The indices are specified as positive or negative numbers based on the normalized angular displacement of the component mean $\Delta\hat{\theta}_k$ relative to the overall distribution mean. For example $k = 1$ will have the lowest positive angular displacement and $k = -1$ has the greatest negative angular displacement. In the case that an odd number of components is used, a $k = 0$ component will be present centered on the mean. This component has a normalized angular displacement of $\Delta\hat{\theta}_0 = 0$ and correlation coefficent $\rho_{r\theta,0} = 0$ for all GM fits due to the symmetry of the distribution. Therefore, these parameters are omitted from the tables to save space.

In the $N_G = 2$ case shown in Table 2, the weights of the two components are equal ($\omega_{-1} = \omega_1 = 0.5$) and the mean angle displacement of the two components is symmetric ($\Delta\hat{\theta}_{-1} = -\Delta\hat{\theta}_1$). The covariance correlation coefficients $\rho_{r\theta,-1} = -\rho_{r\theta,-1}$ are also symmetric, but with the opposite sign convention due to the fact that for a positive angle shift of the mean round the circle the covariance axes are tilted by a negative angle. All other parameters ($\Delta\hat{r}_k$, $\hat{\sigma}_{r,k}$, and $\hat{\sigma}_{\theta,k}$) are identical for the two components. Therefore, the parameters for $k = 1$ and $k = -1$ can be compressed into

one table, with the notation $(\cdot)_{\pm 1}$ for the index expression. Where the parameters differ in sign between $k = 1$ and $k = -1$, $\pm$ or $\mp$ notation is used in conjunction with the parameter value to show this distinction.

Table 2: KL Divergence and Component $k = \pm 1$ Parameters for $N_G = 2$ (sample)

C_B	KL	$\omega_{\pm 1}$	$\Delta\hat{r}_{\pm 1}$	$\Delta\hat{\theta}_{\pm 1}$	$\hat{\sigma}_{r,\pm 1}$	$\hat{\sigma}_{\theta,\pm 1}$	$\rho_{r\theta,\pm 1}$
0.20	0.004	0.500	-0.75	± 0.50	0.95	0.87	∓ 0.21
0.24	0.006	0.500	-0.73	± 0.52	0.94	0.85	∓ 0.22
0.29	0.010	0.500	-0.70	± 0.55	0.92	0.84	∓ 0.23
0.34	0.016	0.500	-0.68	± 0.57	0.89	0.82	∓ 0.24
0.41	0.024	0.500	-0.65	± 0.59	0.86	0.81	∓ 0.25
0.49	0.038	0.500	-0.62	± 0.62	0.82	0.79	∓ 0.26
0.59	0.057	0.500	-0.59	± 0.64	0.78	0.77	∓ 0.27
0.71	0.084	0.500	-0.55	± 0.67	0.74	0.74	∓ 0.27

Tables 3 and 4 show the same parameters calculated for $N_G = 3$. Since this GM has an odd number of components, there exists a $k = 0$ component at zero angular displacement. The parameters for this component are given in Table 3. For components $k = 1$ and $k = -1$ the parameters are given in the single Table 4 due to symmetry. From these tables it can be seen that the general trend in the optimal solutions for $N_G = 3$ is to assign less weight to the central component ω_0 and to further spread the means of the outer components in angular displacement $\Delta\hat{\theta}_{\pm 1}$ as the bias significance increases.

It is evident that these tables with bias significance abscissa are of a small size compared to the original four-dimensional input parameter space of polar mean and covariance. Furthermore, the parameters in the tables change slowly with increasing bias significance (a few percent per line), which means the tables should be robust to interpolation. A full set of tables is available in Appendix A.

Table 3: KL Divergence and Component $k = 0$ Parameters for $N_G = 3$ (sample)

C_B	KL	ω_0	$\Delta\hat{r}_0$	$\hat{\sigma}_{r,0}$	$\hat{\sigma}_{\theta,0}$
0.50	0.009	0.571	-0.47	0.77	0.68
0.57	0.012	0.570	-0.44	0.74	0.66
0.65	0.017	0.566	-0.41	0.69	0.64
0.73	0.024	0.562	-0.38	0.65	0.62
0.83	0.032	0.557	-0.36	0.60	0.60
0.95	0.044	0.553	-0.34	0.56	0.58
1.08	0.058	0.547	-0.32	0.51	0.56
1.22	0.075	0.543	-0.29	0.47	0.54

Table 4: Component $k = \pm 1$ Parameters for $N_G = 3$ (sample)

C_B	$\omega_{\pm 1}$	$\Delta\hat{r}_{\pm 1}$	$\Delta\hat{\theta}_{\pm 1}$	$\hat{\sigma}_{r,\pm 1}$	$\hat{\sigma}_{\theta,\pm 1}$	$\rho_{r\theta,\pm 1}$
0.50	0.214	-0.49	± 1.11	0.81	0.71	∓ 0.31
0.57	0.215	-0.47	± 1.13	0.78	0.69	∓ 0.31
0.65	0.217	-0.45	± 1.15	0.75	0.67	∓ 0.31
0.73	0.219	-0.43	± 1.17	0.71	0.66	∓ 0.31
0.83	0.222	-0.41	± 1.18	0.67	0.64	∓ 0.32
0.95	0.224	-0.38	± 1.20	0.64	0.62	∓ 0.32
1.08	0.226	-0.36	± 1.21	0.60	0.60	∓ 0.33
1.22	0.228	-0.34	± 1.22	0.56	0.59	∓ 0.34

2.3.5 Generating a Mixture From Stored Lookup Table

In order to generate a GM representation for a given radar measurement with range mean $\bar{r}$ and variance σ_r^2 and angle mean $\bar{\theta}$ and variance σ_θ from a lookup table, apply the following steps.

First, evaluate the measurement bias significance C_B using (1). Second, examine the lookup tables for various N_G at this C_B to determine the KL divergence that is achieved and select N_G to limit this value. Finally, retrieve the weights and normalized mean and covariance parameters ω_k, $\Delta\hat{r}_k$, $\Delta\hat{\theta}_k$, $\hat{\sigma}_{M,k}$, $\hat{\sigma}_{N,k}$ for this C_B point in the N_G lookup table (using interpolation if necessary) and apply the following transforms to yield the final mixture means μ_k and covariances Σ_k, where S is given by (57).

$$\mu_k = (\bar{r} + C_B \sigma_r \Delta\hat{r}_k) \begin{bmatrix} \sin(\bar{\theta} + \sigma_\theta \Delta\hat{\theta}_k) \\ \cos(\bar{\theta} + \sigma_\theta \Delta\hat{\theta}_k) \end{bmatrix} \tag{61}$$

$$\Sigma_k = S \begin{bmatrix} \hat{\sigma}_{r,k} & 0 \\ 0 & \hat{\sigma}_{\theta,k} \end{bmatrix} \begin{bmatrix} 1 & \rho_{r\theta,k} \\ \rho_{r\theta,k} & 1 \end{bmatrix} \begin{bmatrix} \hat{\sigma}_{r,k} & 0 \\ 0 & \hat{\sigma}_{\theta,k} \end{bmatrix} S^T \tag{62}$$

2.3.6 Definition of Mixture Lookup Operation

In order to effectively describe the use of the mixture generation procedure in a Kalman filter, it is useful to define a notation which describes it. Let a mixture of size N_G generated using this table lookup procedure for a polar measurement with mean $z = \begin{bmatrix} \bar{r} & \bar{\theta} \end{bmatrix}^T$ and covariance $P_z = \begin{bmatrix} \sigma_r^2 & 0 \\ 0 & \sigma_\theta^2 \end{bmatrix}$ be denoted by

$$\{\omega_k, \mu_k, \Sigma_k\}_{k=1}^{N_G} = \mathrm{Mix}_{\mathrm{EM}}(z, P_z, N_G) \tag{63}$$

and the KL divergence of the mixture predicted by the lookup table be $\mathrm{KL}_{\mathrm{EM}}(z, P_z, N_G)$. Furthermore, suppose a target KL divergence of $KL^\star$ is specified. Then the process to determine the minimum number of components needed to satisfy this limit is given

by

$$N_G^\star(z, P_z, KL^\star) = \arg\min_{N_G} \mathrm{KL}_{\mathrm{EM}}(z, P_z, N_G) \leq KL^\star \tag{64}$$

For convenience, let

$$\mathrm{Mix}_{\mathrm{EM}}^\star(z, P_z, KL^\star) = \mathrm{Mix}_{\mathrm{EM}}(z, P_z, N_G^\star(z, P_z, KL^\star)) \tag{65}$$

and

$$\mathrm{KL}_{\mathrm{EM}}^\star(z, P_z, KL^\star) = \mathrm{KL}_{\mathrm{EM}}(z, P_z, N_G^\star(z, P_z, KL^\star)) \tag{66}$$

be the mixture and evaluated KL divergence which satisfies this limit.

2.3.7 Demonstration of Mixture Modeling Accuracy

In order to demonstrate the effectiveness of this measurement modeling approach, Monte-Carlo runs of the lookup-table-based mixture generation process were performed over a logarithmic grid of angle standard deviations from 1-300 mrad and bias significances from 0.3-10 (which in turn determine the range standard deviation). A fixed range of 100 km was used. The KL divergence from the lookup table (predicted) is compared to the empirical KL divergence calculated by sampling the original polar measurement distribution with 1 million points, converting these points to Cartesian coordinates, and calculating the expectation of the log likelihood.

Figure 4 shows the result of this analysis for $N_G = 3$ components. The upper axes shows the predicted KL divergence for the GM model of the measurement based on the lookup table. The bottom axes shows the KL divergence actually achieved by the samples of the original distribution. It is seen that the predicted KL divergence is matched well in practice by the lookup-table measurement model over the domain of interest. Some slight deviation on the order of 0.01 is observed in the lower bias significance portions of the grid. These deviations are partly due to sampling error as well as the fact that KL divergence is plotted on a logarithmic scale. However, based on Table 1, an $N_G = 3$ model is unlikely to be used outside of $C_B \in \{0.6, 1.1\}$. In this region, the KL divergence is less than 0.06.

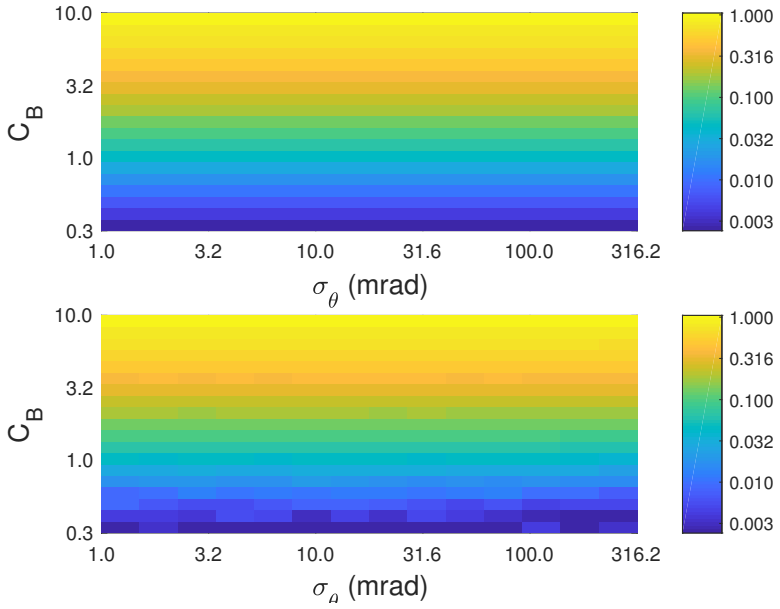

Figure 4: KL Divergence of Generated Mixtures for $N_G = 3$, Predicted (top) and Evaluated (bottom)

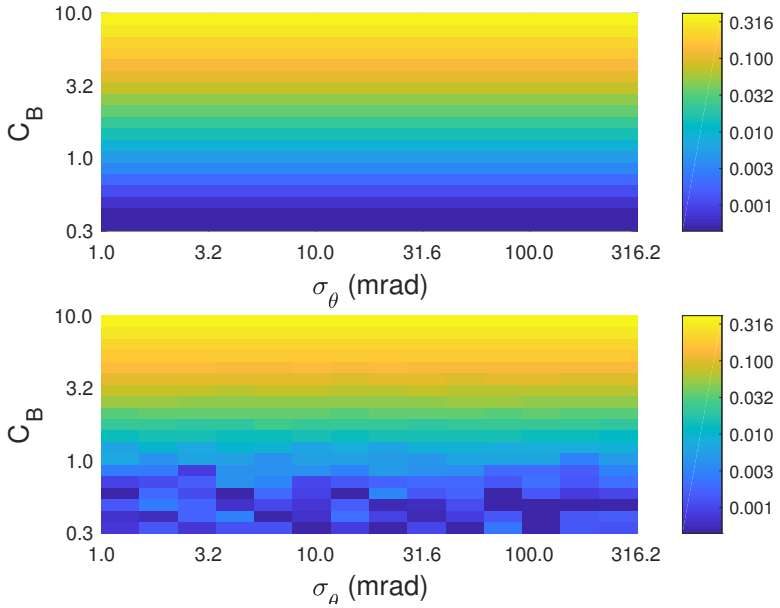

Figure 5: KL Divergence of Generated Mixtures for $N_G = 5$, Predicted (top) and Evaluated (bottom)

Figure 5 shows results for $N_G = 5$ components. Again, the top axes shows the KL divergence predicted by the lookup table while the bottom axes show the KL divergence actually achieved by the model. The KL divergence achieved here is lower (as can be seen by comparison to Figure 4), so sampling error effects can be seen at a higher C_B than in Figure 4 (note also the differing plot scale for the colors). However, again referencing Table 1, an $N_G = 5$ model is most likely to be used for $C_B \in \{0.8, 2.7\}$, where the two charts agree well (differences in the third decimal place).

2.4 Localized Mixture Representations of PDFs

As measurements accumulate in a Kalman filter, the state covariance converges to a level determined by the process noise, measurement rate, and the measurement accuracy. When process noise is low, a significant level of convergence (or variance reduction) will take place. Therefore, the steady-state uncertainty region of the track state may be many times smaller than the uncertainty of the measurement.

In the test cases studied in this chapter, the process noise is small and equal in all spatial dimensions. Therefore, a larger variance reduction is possible in the cross-range dimension than in the range dimension. Furthermore, in several of the filter models studied, the state is represented by a Gaussian mixture, and the extent of the cross-range variance of each individual state component may be many times smaller than the full extent of the state cross-range variance.

In these cases where the state or state component has a much smaller cross-range variance than the measurement, using a KL optimal Cartesian GM model for the measurement will not necessarily give good performance. Recall that the KL divergence is equivalent to the average log-likelihood ratio of the mixture and true PDFs with expectation *taken over the true measurement PDF*. Therefore, an ML fit will optimize the mixture models in order to achieve a good fit at the most likely points

in the measurement PDF. Now, consider the Bayesian update which is performed by the Kalman filter

$$p(x|Z_k) \propto p(z_k|x)p(x|Z_{k-1}) \tag{67}$$

where $Z_k = \{z_k, z_{k-1}, ...\}$ denotes the full set of measurements up to time k and x denotes the state. The Gaussian mixture distribution generated by Mix_{EM} (63) is an approximation of $p(z_k|x)$ optimized over the support of the total measurement distribution. However, with respect to the more limited support of the prior $p(x|Z_{k-1})$, the mixture generated by Mix_{EM} may have reduced performance. This can be seen by examining the definition of the KL divergence.

$$D_{KL}(P_z||Q_z) = \int_{-\infty}^{\infty} p_z(x) \ln \frac{p_z(x)}{q_z(x)}\, dx$$

where $q_z(x)$ is the model PDF and $p_z(x)$ is the true measurement PDF. For simplicity, assume these PDFs are expressed in the same coordinates as the state PDF. Then if the expectation of the the log-likelihood ratio is taken over the prior according to

$$\int_{-\infty}^{\infty} p(x|Z_{k-1}) \ln \frac{p_z(x)}{q_z(x)}\, dx$$

the model $q_z(x)$ which is optimized in terms of KL divergence may have reduced performance.

Figures 6 through 10 illustrate this situation for a settled track component with cross-range support a fraction of that of the measurement. Figure 6 shows a notional Gaussian prior PDF. This PDF is updated by the measurement PDF in Figure 7 to yield the true posterior PDF in Figure 8, which is highly non-Gaussian. (The intensity of the colors in these figures shows the relative density of the PDF, and the colors are thresholded to white at 5 percent of the maximum density.)

Figure 9 shows a six-component GM model of the measurement generate using the lookup table. The two-sigma (Mahalanobis distance = 4) covariance ellipse of each mixture component is plotted as a red line, and the component mean as a

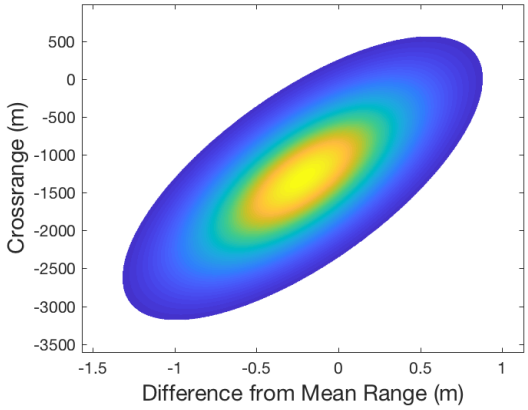

Figure 6: PDF of Prior State Component

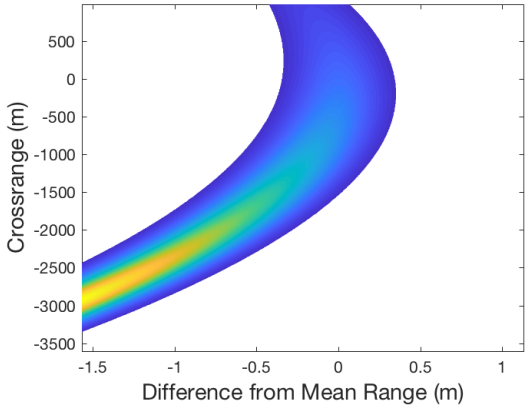

Figure 7: PDF of True Measurement

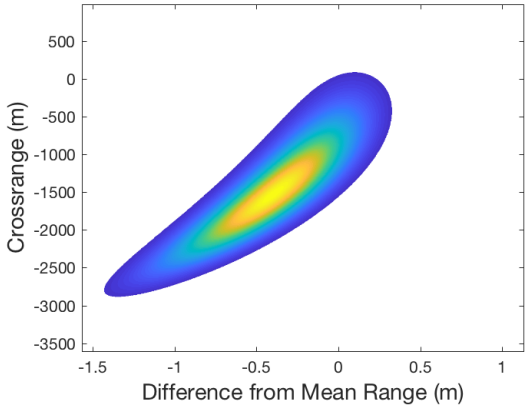

Figure 8: PDF of True Posterior State

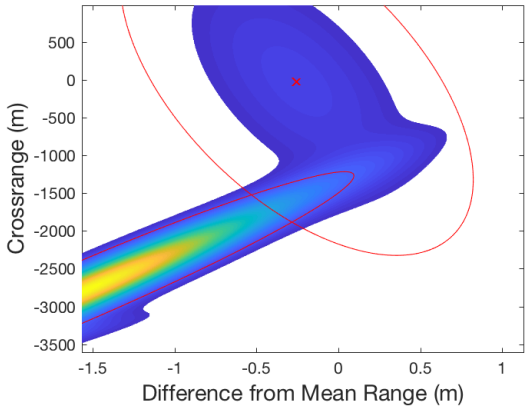

Figure 9: PDF of Six-Component Mixture Measurement

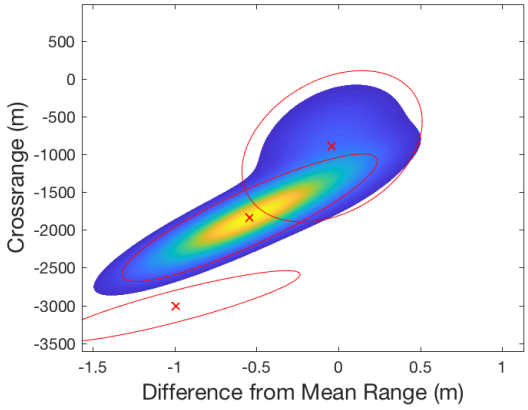

Figure 10: PDF of Six-Component Mixture Posterior State

red cross. Most of the components are not visible due to the plot scale being fixed to the state prior PDF. Note that the PDF density coloring does not fill the top covariance ellipse as this component has very low weight. This model represents the true mixture PDF in Figure 7 well when considered over the entire cross-range of the measurement, but when only the support of the prior state is considered, the mixture is "jagged". This results in a degraded posterior mixture estimate shown in Figure 10 compared to the true posterior in Figure 8. Additionally, most of the six mixture components in the posterior have extremely low weight due to the original measurement components being far from the state. This results in them being unfilled with color representative of the PDF density as the entire component's density is below 5 percent of the maximum. Ideally, these components would be located in a region of higher posterior likelihood, resulting in a more accurate model of the posterior PDF.

In this section, an novel procedure is presented which addresses this difficulty by pre-conditioning the measurement strategically. In order to describe this procedure, the measurement functions to be used in the Kalman filter are defined along with the process for removing measurements from a filtered state by inverting the Bayesian update equations. These concepts will be applied to solve the difficulties associated with the measurement-to-state crossrange covariance mismatch.

2.4.1 Measurement Model Definition

All parametric (non-particle) filters to be considered in this chapter use one of two measurement update methods. Either the state (or state component) is transformed to polar space and the update is made using EKF equations based on the Jacobian of this transformation, or the measurement is transformed via a separate procedure to a set of Cartesian space means and covariances, and the update is made using the linear Kalman filter equations. These two cases are expressed as follows. The

Cartesian state to polar measurement transformation is given by

$$h_P(x) = \begin{bmatrix} \sqrt{x_1^2 + x_2^2} \\ \tan^{-1}(\frac{x_1}{x_2}) \end{bmatrix} \tag{68}$$

$$H_P(x) = \frac{\partial}{\partial x} h_P(x) = \begin{bmatrix} \frac{x_1}{\sqrt{x_1^2+x_2^2}} & \frac{x_2}{\sqrt{x_1^2+x_2^2}} & 0 & 0 \\ \frac{x_2}{x_1^2+x_2^2} & -\frac{x_1}{x_1^2+x_2^2} & 0 & 0 \end{bmatrix} \tag{69}$$

The inverse of this function (with respect to the Cartesian position only) is also occasionally needed and is expressed by

$$h_P^{-1}(z) = z_1 \begin{bmatrix} \sin z_2 \\ \cos z_2 \end{bmatrix} \tag{70}$$

where z_1 and z_2 denote the components of the vector z. Additionally, sometimes only the angle coordinate is of interest. Let this transformation to a scalar be expressed by

$$h_\theta(x) = \tan^{-1}\left(\frac{x_1}{x_2}\right) \tag{71}$$

$$H_\theta(x) = \frac{\partial}{\partial x} h_\theta(x) = \begin{bmatrix} \frac{x_2}{x_1^2+x_2^2} & -\frac{x_1}{x_1^2+x_2^2} & 0 & 0 \end{bmatrix} \tag{72}$$

If the measurements are converted to Cartesian space, then the observation is simply a linear projection onto the position subspace as given by

$$h_C(x) = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = H_C x \tag{73}$$

where

$$H_C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \tag{74}$$

2.4.2 Kalman Update Parameterization

For convenience, let the standard Kalman filter update equations be expressed as a set of functions

$$K_{(\cdot)}(x, P, z, R) = PH_{(\cdot)}^T(x)(H_{(\cdot)}(x)PH_{(\cdot)}^T(x) + R)^{-1} \tag{75}$$

$$U_{x,(\cdot)}(x, P, z, R) = x + K_{(\cdot)}(z - h_{(\cdot)}(x)) \tag{76}$$

$$U_{P,(\cdot)}(x, P, z, R) = \left(I - K_{(\cdot)}H_{(\cdot)}(x)\right) P \tag{77}$$

where $(\cdot)$ is a way to generalize the measurement function that is used. For example, if the Cartesian measurement function h_C and linearization matrix H_C were used, then the mean update at time k is given by

$$x_{k|k} = U_{x_{k|k-1},C}(x_{k|k-1}, P_{k|k-1}, z_k, R_k) = x_{k|k-1} + K_k(z_k - h_C(x_{k|k-1}))$$

where

$$\begin{aligned} K_k &= K_C(x_{k|k-1}, P_{k|k-1}, z_k, R_k) \\ &= P_{k|k-1}H_C^T(x_{k|k-1})(H_C(x_{k|k-1})P_{k|k-1}H_C^T(x_{k|k-1}) + R_k)^{-1} \end{aligned}$$

The use of the $(\cdot)$ notation therefore avoids passing the measurement model as an extra function argument.

Furthermore, define the Mahalanobis distance associated with the measurement log likelihood under the given state as

$$MD_{z,(\cdot)}(x, P, z, R) = \left(z - h_{(\cdot)}(x)\right)^T S_{z,(\cdot)}(x, P, z, R)^{-1} \left(z - h_{(\cdot)}(x)\right) \tag{78}$$

where

$$S_{z,(\cdot)}(x, P, z, R) = H_{(\cdot)}(x)PH_{(\cdot)}^T(x) + R \tag{79}$$

The full likelihood function associated with this is

$$L_{z,(\cdot)}(x, P, z, R) = \frac{1}{\sqrt{\left|2\pi S_{z,(\cdot)}(x, P, z, R)\right|}} \exp\left(-\frac{1}{2} MD_{z,(\cdot)}(x, P, z, R)\right) \tag{80}$$

2.4.3 Measurement Inverse Filtering

Let $\hat{x}_{k|k-1}$ and $P_{k|k-1}$ represent the current prior state mean and covariance, and z_k and R_k represent the polar measurement mean and covariance. Now, note that the standard EKF update equations

$$K = P_{k|k-1}H^T \left(HP_{k|k-1}H^T + R_k\right)^{-1} \tag{81}$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K(z_k - h(\hat{x}_{k|k-1})) \tag{82}$$

$$P_{k|k} = (I - KH)\, P_{k|k-1} \tag{83}$$

may be inverted to retrieve an approximation of the prior state and covariance as follows. Given

$$K = P_{k|k} H^T R_k^{-1} \tag{84}$$

$$K' = P_{k|k} H^T \left(R_k - H P_{k|k} H^T\right)^{-1} \tag{85}$$

then

$$\hat{x}_{k|k-1} = (I + K'H)\, \left(\hat{x}_{k|k} - K(z_k - h(\hat{x}_{k|k}) + H\hat{x}_{k|k})\right) \tag{86}$$

$$P_{k|k-1} = (I + K'H)\, P_{k|k} \tag{87}$$

where H is evaluated at the posterior. If the nonlinearity of the function h is not too severe, and $\hat{x}_{k|k}$ and $\hat{x}_{k|k-1}$ are "close" relative to this non-linearity, then these equations will give a reasonable estimate of the prior. For a linear measurement function ($h(x) = Hx$), these equations give an exact solution. In the general case, the inverse in (85) is not guaranteed to exist. However, this equation is derived by rearranging the information Kalman filter update and applying the Woodbury matrix inverse formula. If the information being removed from the posterior is not large compared to the total information, the matrix should remain invertible (provided H is a reasonable linear approximation of the update). Since the conditioning measurement is constructed by design with less information in cross-range (angle) than the state prior, this will be the case.

Let this process be represented by the functions

$$K'_{(\cdot)}(x, P, z, R) = P H^T_{(\cdot)}(x) \left(R - H_{(\cdot)}(x) P H^T_{(\cdot)}(x)\right)^{-1} \tag{88}$$

$$\hat{K}_{(\cdot)}(x, P, z, R) = P H^T_{(\cdot)}(x) R^{-1} \tag{89}$$

Then the inverse filtering of the mean is given by

$$U^{-1}_{x,(\cdot)}(x, P, z, R) = \left(I + K'_{(\cdot)} H_{(\cdot)}(x)\right)\, \left(x - \hat{K}_{(\cdot)}(z - h(x) + H_{(\cdot)}(x)x)\right) \tag{90}$$

The inverse filtering of the covariance is given by

$$U^{-1}_{P,(\cdot)}(x, P, z, R) = \left(I + K'_{(\cdot)} H_{(\cdot)}(x)\right) P \tag{91}$$

where again the $(\cdot)$ notation is used to generalize the measurement model (see Section 2.4.1), and the arguments to $\hat{K}_{(\cdot)}$ and $K'_{(\cdot)}$ have been suppressed for brevity.

2.4.4 Solution to the Measurement Mixture and State Support Mismatch

This inverse filtering process is applied to solve the problem of support mismatch in the state and measurement mixtures with the following five steps.

1. Generate a fake "conditioning" angle-only measurement with strategically chosen mean and covariance. The mean and covariance are chosen such that the following step will move the measurement mean on top of the state mean in cross-range and reduce the measurement cross-range variance to be comparable to the cross-range variance of the state.

2. Update the original measurement in polar space by conditioning it with the fake measurement. The effect of this conditioning changes the angle (cross-range) support of the modified measurement to cover the state cross-range support more tightly.

3. Generate a Gaussian mixture to model the conditioned measurement in Cartesian space using $\text{Mix}^{\star}_{\text{EM}}$.

4. Update the prior state to a posterior mixture using the conditioned measurement mixture in Cartesian space.

5. Remove the effect of the conditioning measurement on the posterior state using the inverse filtering process as if it had been applied independently to the state using EKF.

Let the measurement angle at time k along with its variance be given by $\theta_{z,k}$ and $\sigma^2_{\theta,z,k}$. Then the prior state angular mean and variance are given by $\theta_{x,k} = h_\theta(\hat{x}_{k|k-1})$ and $\sigma^2_{\theta,x,k} = H_\theta(\hat{x}_{k|k-1})P_{k|k-1}H_\theta(\hat{x}_{k|k-1})^T$. The goal now is to generate a conditioning measurement $\theta_{y,k}, \sigma^2_{\theta,y,k}$ such that if $\theta_{z,k}, \sigma^2_{\theta,z,k}$ is updated by this measurement using Bayesian Gaussian conditioning (Kalman filter update), the resulting modified measurement has a mean and covariance

$$\theta_{z',k} = U_{x,I}(\theta_{z,k}, \sigma^2_{\theta,z,k}, \theta_{y,k}, \sigma^2_{\theta,y,k}) = \theta_{x,k} \tag{92}$$

and

$$\sigma^2_{\theta,z',k} = U_{P,I}(\theta_{z,k}, \sigma^2_{\theta,z,k}, \theta_{y,k}, \sigma^2_{\theta,y,k}) = \kappa\sigma^2_{\theta,x,k} \tag{93}$$

where $\kappa \geq 1$ is a "covering factor" used to ensure the conditioned measurement covers the state cross-range support sufficiently, and $h_I(x) = x$, $H_I = I$ are identity measurement functions describing the model used in the update $U_{x,I}, U_{P,I}$.

For the scalar case, it is straightforward to show that $\theta_{y,k}, \sigma^2_{\theta,y,k}$ are given by

$$\theta_{y,k} = \frac{\sigma^2_{\theta,z,k}\theta_{x,k} - \kappa\sigma^2_{\theta,x,k}\theta_{z,k}}{\sigma^2_{\theta,z,k} - \kappa\sigma^2_{\theta,x,k}} \tag{94}$$

$$\sigma^2_{\theta,y,k} = \frac{\kappa\sigma^2_{\theta,x,k}\sigma^2_{\theta,z,k}}{\sigma^2_{\theta,z,k} - \kappa\sigma^2_{\theta,x,k}} \tag{95}$$

If the measurement covariance is decoupled in range and angle (as in the cases modeled by Mix_{EM}) and $\theta_{y,k}, \sigma^2_{\theta,y,k}$ are generated from (94) and (95), then the angular mean and covariance of the conditioned measurement z'_k and R'_k are equal to $\theta_{x,k}$ and $\kappa\sigma^2_{\theta,x,k}$ respectively. Note that if $\sigma^2_{\theta,z,k} < \kappa\sigma^2_{\theta,x,k}$, then the cross-range support of the measurement is already well matched to the state. In this case, skip the conditioning step and keep $z'_k = z_k$ and $R'_k = R_k$.

Once the conditioned measurement z'_k and R'_k has been created, the lookup table is used to generate a GM model of this measurement $\left\{\omega'_{k,z_C,j}, z'_{k,C,j}, R'_{k,C,j}\right\}_{j=1}^{N^\star_G} = \text{Mix}^\star_{\text{EM}}(z'_k, R'_k, KL^\star)$. For each component j in this GM, apply a Kalman update to

the prior state, resulting in "conditioned" posterior state components

$$\hat{x}'_{k|k,j} = U_{x,C}(\hat{x}_{k|k-1}, P_{k|k-1}, z'_{k,C,j}, R'_{k,C,j}) \tag{96}$$

$$P'_{k|k,j} = U_{P,C}(\hat{x}_{k|k-1}, P_{k|k-1}, z'_{k,C,j}, R'_{k,C,j}) \tag{97}$$

with weights $\omega'_{k,x,j} = \omega'_{k,z_C,j}$.

Given this conditioned posterior GM, the final step is to apply the inverse filtering process to remove the conditioning effect of $\theta_{y,k}$ and $\sigma^2_{\theta,y,k}$. For the mean and covariance, calculate

$$\hat{x}_{k|k,j} = U^{-1}_{x,\theta}(\hat{x}'_{k|k,j}, P'_{k|k,j}, \theta_{y,k}, \sigma^2_{\theta,y,k}) \tag{98}$$

$$P_{k|k,j} = U^{-1}_{P,\theta}(\hat{x}'_{k|k,j}, P'_{k|k,j}, \theta_{y,k}, \sigma^2_{\theta,y,k}) \tag{99}$$

To adjust the weights, note that if the true posterior weights are $\omega_{k,x,j}$, then a forward update of a mixture filter applying the measurement $\theta_{y,k}$ and $\sigma^2_{\theta,y,k}$ is given by

$$\omega'_{k,x,j} = \frac{\omega_{k,x,j} L_{z,\theta}\left(\hat{x}_{k|k,j}, P_{k|k,j}, \theta_{y,k}, \sigma^2_{\theta,y,k}\right)}{\sum_{j=1}^{N^\star_G} \omega_{k,x,j} L_{z,\theta}\left(\hat{x}_{k|k,j}, P_{k|k,j}, \theta_{y,k}, \sigma^2_{\theta,y,k}\right)} \tag{100}$$

which can be inverted to yield

$$\omega_{k,x,j} = \frac{\frac{\omega'_{k,x,j}}{L_{z,\theta}\left(\hat{x}_{k|k,j}, P_{k|k,j}, \theta_{y,k}, \sigma^2_{\theta,y,k}\right)}}{\sum_{j=1}^{N^\star_G} \frac{\omega'_{k,x,j}}{L_{z,\theta}\left(\hat{x}_{k|k,j}, P_{k|k,j}, \theta_{y,k}, \sigma^2_{\theta,y,k}\right)}} \tag{101}$$

Note that the likelihoods are evaluated at the final version of the posterior GM means and covariances, not the conditioned version. Then the mixture $\left\{\omega_{k,x,j}, \hat{x}_{k|k,j}, P_{k|k,j}\right\}_{j=1}^{N^\star_G}$ represents the posterior PDF estimate of the prior state $\hat{x}_{k|k-1}$ and $P_{k|k-1}$ updated by z_k and R_k.

The steps of this process are illustrated in Figures 11 through 13. Figure 11 shows the conditioned measurement mixture $\left\{\omega'_{k,z_C,j}, z'_{k,C,j}, R'_{k,C,j}\right\}$ when $N^\star_G = 6$. A value of $\kappa = 1.5$ was used to define the conditioned measurement angle support. In contrast to the original unconditioned mixture in Figure 9, the six components are all located near the the support of the state prior in Figure 6. Applying the Kalman update to

the prior yields the conditioned posterior $\left\{\omega'_{k,x,j}, \hat{x}'_{k|k,j}, P'_{k|k,j}\right\}$ in Figure 12. This is close to the true posterior in Figure 8, but the spread of the PDF is slightly smaller in cross-range. After applying the inverse filtering process to remove the preconditioning in angle, the posterior mixture in Figure 13 is achieved, which is an excellent match to the true posterior.

This process was repeated for a lower-fidelity mixture model with $N_G^\star = 3$ components, and the results are shown in Figures 14 through 16. Even with this lower number of mixture components, the accuracy of the final posterior PDF is much greater than in the case of Figure 10, which used a six-component measurement mixture model without conditioning.

Note that in the case that the state prior is represented by a mixture, this process is applied to each state component individually to achieve the mixture PDF of the posterior under the condition that the current component is the true component. To find the posterior PDF of the unconditioned state, the individual posterior mixture weights must be combined with the prior weights taking into account the likelihood of the prior components under the measurement. This process is described in further detail in the full filter specification.

2.5 Application of Gaussian Mixture Contact-Lens Model to Radar Tracking

It now remains to evaluate the effectiveness of this measurement modeling technique when applied to a GM Kalman filter. In order to provide an efficient implementation, ideas from several previously proposed filters (MCAEKF [27] and CbGMF [28]) are combined with the Gaussian mixture measurement model derived above. To control the number of Gaussian mixtures in the track representation, a variation of the clustering algorithm proposed by Salmond [25] is used. This filter is called the Measurement-Adaptive Gaussian Mixture Filter (MAGMF). The performance of this

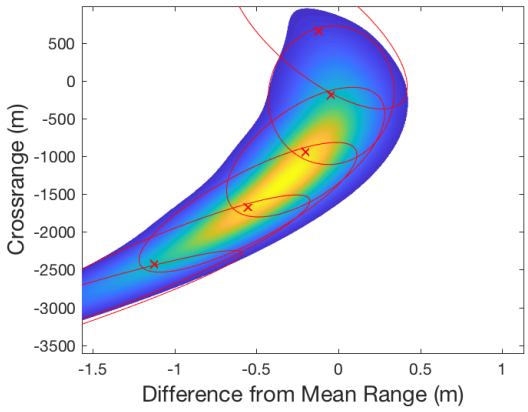

Figure 11: PDF of Six-Component $\left\{\omega'_{k,z_C,j}, z'_{k,C,j}, R'_{k,C,j}\right\}$ Conditioned Mixture Measurement

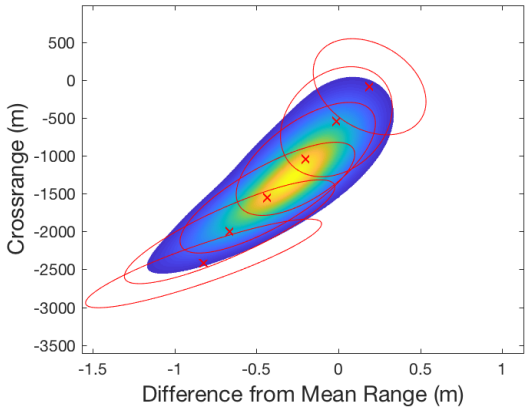

Figure 12: PDF of Six-Component $\left\{\omega'_{k,x,j}, \hat{x}'_{k|k,j}, P'_{k|k,j}\right\}$ Conditioned Mixture Posterior State

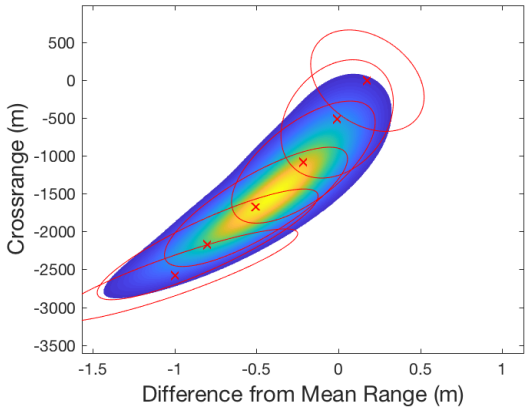

Figure 13: PDF of Six-Component $\left\{\omega_{k,x,j}, \hat{x}_{k|k,j}, P_{k|k,j}\right\}$ Inverse Filtered Mixture Posterior State

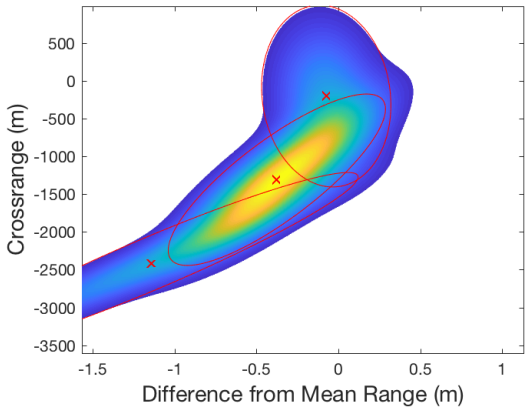

Figure 14: PDF of Three-Component $\left\{\omega'_{k,z_C,j}, z'_{k,C,j}, R'_{k,C,j}\right\}$ Conditioned Mixture Measurement

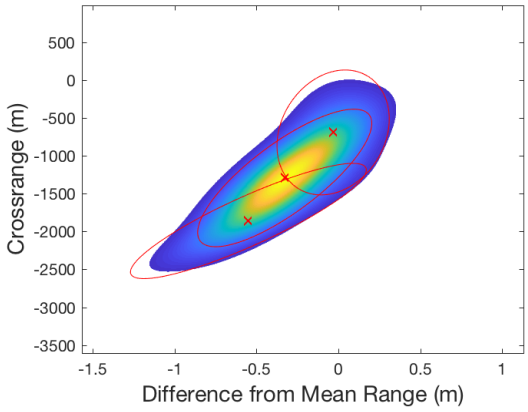

Figure 15: PDF of Three-Component $\left\{\omega'_{k,x,j}, \hat{x}'_{k|k,j}, P'_{k|k,j}\right\}$ Conditioned Mixture Posterior

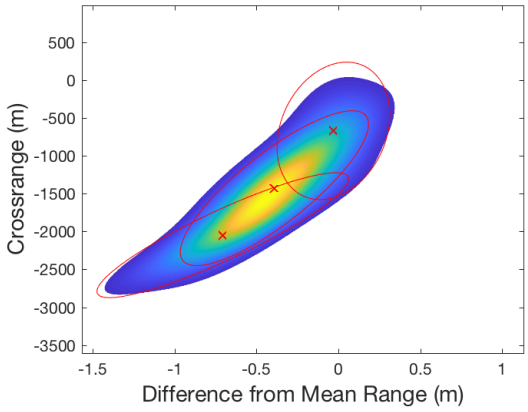

Figure 16: PDF of Three-Component $\left\{\omega_{k,x,j}, \hat{x}_{k|k,j}, P_{k|k,j}\right\}$ Inverse Filtered Mixture Posterior

new filter is compared to a variety of filtering techniques including

1. Standard EKF with Cartesian state and polar measurement
2. Converted measurement EKF without bias compensation
3. Converted measurement EKF with bias compensation [19]
4. UKF with Cartesian state and polar measurement [16]
5. MCAEKF [27]
6. CbGMF [28]
7. GMM-ITS [32]
8. Regularized Particle Filter [24]

2.5.1 Filter Dynamics Model

For each filter studied, the dynamics model used is a nearly-constant-velocity (NCV) with small process noise $Q_c = 0.001I_2$ (where I_2 is an identity matrix in 2D). This linear model has the following mean and covariance update equations which can be found in standard references [2]. Let

$$\hat{x}_k = \begin{bmatrix} \hat{x}_{1,k} \\ \hat{x}_{2,k} \\ \hat{\dot{x}}_{1,k} \\ \hat{\dot{x}}_{2,k} \end{bmatrix} \tag{102}$$

$$F(\Delta t) = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{103}$$

$$Q(\Delta t) = \begin{bmatrix} \frac{1}{3}Q_c\Delta t^3 & \frac{1}{2}Q_c\Delta t^2 \\ \frac{1}{2}Q_c\Delta t^2 & Q_c\Delta t \end{bmatrix} \tag{104}$$

where $\hat{x}_{1,k}$ and $\hat{x}_{2,k}$ represent the scalar position components of the state vector and $\hat{\dot{x}}_{1,k}$ and $\hat{\dot{x}}_{2,k}$ represent their corresponding scalar velocities. Then the prediction equations are given by

$$\hat{x}_{k|k-1} = F(\Delta t)\hat{x}_{k-1|k-1} \tag{105}$$

$$P_{k|k-1} = F(\Delta t)P_{k-1|k-1}F(\Delta t)^T + Q(\Delta t) \tag{106}$$

2.5.2 Filter Initialization

All non-particle filters are initialized using the so-called "one-point" initialization technique. For single-Gaussian filters, the position portion of $\hat{x}_{0|0}$ and $P_{0|0}$ are set to the mean and covariance of the converted measurement z_0 and R_0. For the Gaussian mixture filters (MAGMF, GMM-ITS), a mixture representation of the position distribution $\left\{\omega_{0|0,i}, \hat{x}_{0|0,i}P_{0|0,i}\right\}_{i=1}^{N_G}$ is generated by the respective techniques used in the filter. The velocity portion of the state is initialized to a mean of zero and a diagonal covariance of $\frac{v_{max}^2}{3}$, where $v_{max} = 300$ m/s is a prior knowledge parameter of the maximum speed of the target.

For the RPF, the position components of the initial particles $x_{0|0,i}$ are drawn from the Gaussian measurement distribution in polar space and these samples then converted to Cartesian space by way of the standard transformation (4) and (5). The velocity particles are drawn using the two-point differencing method described by Romeo [24]. (A one-point style initialization was attempted for the RPF, but does not have good performance due to the size of the velocity prior). All initial weights $\omega_{0|0,i}$ are set uniformly to $\frac{1}{N_p}$, where N_p is the number of particles.

This one-point initialization technique works well for a linear dynamics as the large uncertainty in velocity does not affect the linearization of the equations. Furthermore, this method avoids the extra effort associated with curve fitting/batch estimation

frequently used to initialize filters. Alternative techniques may be required in the case of non-linear dynamics.

2.5.3 Unscented Transform

The full Unscented transform (UT) procedure will not be described in detail here, as it is covered in detail in multiple references [16][30]. If $z = h(x)$ is the transformation function from x to z, then if the mean and covariance of x are $\bar{x}$ and P_x, the Unscented transform calculates emperically the mean of z ($\bar{z}$), the covariance of z (P_z), and the cross-covariance between x and z (P_{xz}). Let these be expressed as

$$(\bar{z}, P_z, P_{xz}) = \mathrm{UT}\left(h, \bar{x}, P_x\right) \tag{107}$$

This Unscented transform function is used to calculate the true moments of converted measurements as well as the terms needed for the Kalman gain in the UKF.

2.5.4 Mean and Covariance of a Mixture

Given a (partial) Gaussian mixture PDF described by the parameters

$$\Theta = \{\omega_i, \mu_i, \Sigma_i\}_{i=1}^{N_G}$$

the mean and covariance are given by

$$\mu(\Theta) = \frac{\sum_{i=1}^{N_G} \omega_i \mu_i}{\sum_{i=1}^{N_G} \omega_i} \tag{108}$$

$$\Sigma(\Theta) = \frac{\sum_{i=1}^{N_G} \omega_i \left(\Sigma_i + (\mu_i - \mu(\Theta))(\mu_i - \mu(\Theta))^T\right)}{\sum_{i=1}^{N_G} \omega_i} \tag{109}$$

In the case of a complete GM PDF, the sum of weights $\sum_{i=1}^{N_G} \omega_i$ is one. However, in order to allow the computation of means and covariances of subsets of the components, normalization by the sum of the weights is performed.

A common operation encountered in GM filtering is to combine a particular subset of components into a single component by moment matching. Let this operation be defined as

$$\mathrm{ReduceGM}(\Theta) = \left\{ \sum_{i=1}^{N_G} \omega_i, \mu(\Theta), \Sigma(\Theta) \right\} \tag{110}$$

2.5.5 Description of Alternative Filter Algorithms

The application of the preceding subalgorithms will now be used to describe the alternative filters under consideration.

2.5.5.1 Standard Polar Space EKF

For the polar space EKF, the previous posterior defined by $\hat{x}_{k-1|k-1}$ and $P_{k-1|k-1}$ is propagated using (105) and (106) to the current prior defined by $\hat{x}_{k|k-1}$ and $P_{k|k-1}$. The prior state predicts the measurements in polar space, and the filter is updated using $\hat{x}_{k|k} = U_{x,P}(\hat{x}_{k|k-1}, P_{k|k-1}, z_k, R_k)$ and $P_{k|k} = U_{P,P}(\hat{x}_{k|k-1}, P_{k|k-1}, z_k, R_k)$.

2.5.5.2 Converted Measurements without Bias Compensation

In this case, Cartesian pseudomeaurements $z_{C,k}$ and $R_{C,k}$ are formed using a first-order linearization, i.e. $z_{C,k} = h_P^{-1}(z_k)$ and $R_{C,k} = \left(\frac{\partial}{\partial z} h_P^{-1}\right) R_k \left(\frac{\partial}{\partial z} h_P^{-1}\right)^T$.

The previous posterior $\hat{x}_{k-1|k-1}$ and $P_{k-1|k-1}$ is propagated using (105) and (106) to the current prior $\hat{x}_{k|k-1}$ and $P_{k|k-1}$, and a Cartesian space update using the converted measurements occurs according to $\hat{x}_{k|k} = U_{x,C}(\hat{x}_{k|k-1}, P_{k|k-1}, z_{C,k}, R_{C,k})$ and $P_{k|k} = U_{P,C}(\hat{x}_{k|k-1}, P_{k|k-1}, z_{C,k}, R_{C,k})$.

2.5.5.3 Converted Measurements with Bias Compensation

This case is identical to the case without bias compensation, except that $z_{C,k}, R_{C,k}$ corresponds to the true mean and covariance of the measurement distribution in Cartesian space. This may be evaluated by the methods of [19] and [20] or by Unscented Transform. In this case, the UT method was chosen with

$$(z_{C,k}, R_{C,k}) = \mathrm{UT}\left(h_P^{-1}, z_k, R_k\right) \tag{111}$$

2.5.5.4 UKF in Polar Space

The prior $\hat{x}_{k|k-1}$ and $P_{k|k-1}$ is computed with (105) and (106) as in the other filters. (In a usual UKF, this propagation is done using UT, but since the dynamics are linear

here, there is no benefit to using this). For the update, a variation on the Kalman filter equation is used. First, the mean, covariance, and cross-covariance of the state conversion to the polar measurement are found using UT.

$$\left(\hat{z}_k, \hat{R}_k, P_{xz,k}\right) = \text{UT}\left(h_P, \hat{x}_{k|k-1}, P_{k|k-1}\right) \tag{112}$$

Then

$$K_k = P_{xz,k}(\hat{R}_k + R_k)^{-1} \tag{113}$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k(z_k - \hat{z}_k) \tag{114}$$

$$P_{k|k} = P_{k|k-1} - K_k(\hat{R}_k + R_k)K_k^T \tag{115}$$

2.5.5.5 MCAEKF

MCAEKF [27] proceeds almost identically to the standard Polar EKF. However, at each time step the angle variance of the state $\sigma^2_{\hat{\theta},k} = H_\theta P_{k|k-1} H_\theta^T$ is computed (where H_θ is evaluated at $\hat{x}_{k|k-1}$). The polar bias significance is then computed using

$$C_B = \frac{r_k \sigma^2_{\hat{\theta},k}}{2\sigma_{r,k}} \tag{116}$$

where r_k and $\sigma_{r,k}$ are the range and range standard deviation of the measurement.

If the bias significance exceeds a given threshold $C_B^\star$, then the range standard deviation of the measurement is set to $\sigma'_{r,k} = \frac{C_B}{C_B^\star}\sigma_r$ and substituted into the modified polar covariance R'_k. The update then proceeds normally with

$$\hat{x}_{k|k} = U_{x,P}(\hat{x}_{k|k-1}, P_{k|k-1}, z_k, R'_k)$$

and

$$P_{k|k} = U_{P,P}(\hat{x}_{k|k-1}, P_{k|k-1}, z_k, R'_k)$$

Note that in the original work [27], the covariance adaptation is given for a three-dimensional case in terms of the condition ratio measure. Since the bias significance, condition ratio, and KL Divergence were shown to be equivalent, the inflation rule is expressed here in terms of bias significance. A limit of $C_B^\star = 0.2$ is used based on the reported results in [19].

2.5.5.6 CbGMF

CbGMF [28] is a variation on MCAEKF in which the state PDF is split into a number of components along the crossrange axis so that the bias significance specified by (116) for each given component is limited to $C_B^\star$. Let the current prior state mixture PDF be described by the weights, means, and covariances $\left\{\omega_{k|k-1,i}, \hat{x}_{k|k-1,i}, P_{k|k-1,i}\right\}_{i=1}^{N_G}$. (Each prior component is propagated one-to-one from a previous posterior component $\hat{x}_{k-1|k-1,i}, P_{k-1|k-1,i}$ using (105) and (106) with no modification to the weights)

For each state component i, evaluate the bias significance (116). If it exceeds $C_B^\star$, that component is split into a new Gaussian mixture with $N'_{G,i}$ components $\left\{\omega_{k|k-1,i}, \hat{x}_{k|k-1,i}, P_{k|k-1,i}\right\}_{j=1}^{N'_{G,i}}$ such that the bias significance limit is satisfied. The split procedure is described in [28] and is not reproduced in detail for brevity, but it generally consists of splitting the crossrange support into $N'_{G,i}$ equal intervals and conditioning the original distribution on 1-D Gaussians in cross-range centered on those intervals. The correct number of Gaussians $N'_{G,i}$ is found by trial and error until the desired threshold is met.

Let the prior distribution PDF of the mixture after this splitting has been performed on every component be represented by $\left\{\omega_{k|k-1,i'}, \hat{x}_{k|k-1,i'}, P_{k|k-1,i'}\right\}_{i'=1}^{N'_G}$. (Here the i, j indices have been relabeled to the single index i'.) Then, for all i', update the prior component means and covariances using

$$\hat{x}_{k|k,i'} = U_{x,P}(\hat{x}_{k|k-1,i'}, P_{k|k-1,i'}, z_k, R_k) \tag{117}$$

and

$$P_{k|k,i'} = U_{P,P}(\hat{x}_{k|k-1,i'}, P_{k|k-1,i'}, z_k, R_k) \tag{118}$$

The weights are updated using

$$\omega_{k|k,i'} = \omega_{k|k-1,i'} \frac{L_{z,P}(\hat{x}_{k|k-1,i'}, P_{k|k-1,i'}, z_k, R_k)}{\sum_{i'=1}^{N'_G} L_{z,P}(\hat{x}_{k|k-1,i'}, P_{k|k-1,i'}, z_k, R_k)} \tag{119}$$

where $L_{z,P}$ is the likelihood defined in (80).

The CbGMF implements a component control algorithm to limit the number of Gaussians components that are maintained. First, "diverged" track components are dropped. In the original work, this was done by maintaining a cumulative sum of the Mahalanobis distances $MD_{z,P}(\hat{x}_{k|k-1,i'}, P_{k|k-1,i'}, z_k, R_k)$ over time. However, this proves to be problematic due to the constant splitting of tracks - it is not clear how the distance history should be maintained across a split. For this research components with very low weight were simply dropped.

The second aspect of component control in the CbGMF is that if the number of states exceeds a threshold, all state components are collapsed to a single Gaussian, and the splitting procedure is run anew on the next filter iteration. Depending on the state of convergence of the filter, this may discard significant information in range, as the fine structure is essentially averaged out over crossrange.

2.5.5.7 GMM-ITS

The GMM-ITS filter described in [32, 33] adopts a similar approach to the MAGMF in that the measurement is modeled by a GM. However, the measurement GM is not generated using ML, but using a procedure similar to the splitting of the state in the CbGMF. Given a prior state mixture $\left\{\omega_{x,k|k-1,i}, \hat{x}_{k|k-1,i}, P_{k|k-1,i}\right\}_{i=1}^{N_{G,x}}$, each component of the mixture i is updated in Cartesian space using a Cartesian converted mixture representation of the measurement $\{\omega_{z_C,k,j}, z_{C,k,j}, R_{C,k,j}\}_{j=1}^{N_{G,z}}$. The method used to generate the mixture consists of placing mixture means equally across the angular support with equal mixture covariances. The weights are then defined proportional to the likelihood of the mixture means under the original PDF. Additionally, an inflation factor is multiplied by the original range variance to generate the mixure component range variances. The full procedure is detailed in [32] and omitted here for brevity. The $N_{G,x} \times N_{G,z}$ updates are summarized by

$$\hat{x}_{k|k,i,j} = U_{x,C}(\hat{x}_{k|k-1,i}, P_{k|k-1,i}, z_{C,k,j}, R_{C,k,j}) \tag{120}$$

and

$$P_{k|k,i,j} = U_{P,C}(\hat{x}_{k|k-1,i}, P_{k|k-1,i}, z_{C,k,j}, R_{C,k,j}) \tag{121}$$

The weights are updated using

$$\omega_{k|k,i,j} = \frac{\omega_{k|k-1,i}\omega_{z_C,k,j}L_{z,C}(\hat{x}_{k|k-1,i}, P_{k|k-1,i}, z_{C,k,j}, R_{C,k,j})}{\sum_{i,j}\omega_{k|k-1,i}\omega_{z_C,k,j}L_{z,C}(\hat{x}_{k|k-1,i}, P_{k|k-1,i}, z_{C,k,j}, R_{C,k,j})} \tag{122}$$

where the likelihoods are calculated using the converted measurement distribution.

For component control, the GMM-ITS applies the simple strategy of merging all posterior components associated with a single measurement component j to a single Gaussian. Additionally, components with probability lower than a given threshold are discarded altogether. More details are available in [33].

During the research performed in this chapter, the KL divergence of the GM measurement models used by the GMM-ITS from the true measurement distribution was computed. It was found that due to the range inflation factor applied to the component covariances the KL divergence has a fundamental lower bound which is higher than the divergence achieved by the ML GM models with an appropriate choice of N_G. Therefore, more GM components are required to represent the measurement accurately.

2.5.6 Novel MAGMF Implementation

The new MAGMF uses ML optimal GM representations of the measurements and incorporates ideas from the MCAEKF, CbGMF, and GMM-ITS approaches adaptively to achieve the best performance. The major differences from these filters are as follows.

1. The MAGMF uses the MCAEKF approach to update single track components. However, it only does this when the cost of the inflation to range covariance is low compared to the cost of using a Cartesian mixture.

2. The MAGMF relies on splitting of the track representation to maintain consistency as in the CbGMF. However, the splitting does not occur based on an ad-hoc procedure, but by the natural process of a measurement update. Additionally, the tracks need not be split to the extent that all components satisfy the consistency rule of MCAEKF at all times, since a measurement mixture is used for the update.

3. The MAGMF uses a Cartesian mixture modeling strategy for the measurement as does the GMM-ITS. However, the models are generated using lookup tables based on ML fitting as opposed to the ad-hoc strategy used in [32]. Additionally, a novel procedure is used to ensure that the measurement model components are located primarily within the cross-range support of the track.

The filter is initialized with a mixture by the one-point method described in Section 2.5.2, where the initial position mixture is given by $\text{Mix}^{\star}_{\text{EM}}(z_0, R_0, KL^{\star})$ (65).

For each time step k, the previous posterior $\left\{\omega_{x,k-1|k-1,i}, \hat{x}_{k-1|k-1,i}, P_{k-1|k-1,i}\right\}_i$ is updated to the current prior $\left\{\omega_{x,k|k-1,i}, \hat{x}_{k|k-1,i}, P_{k|k-1,i}\right\}_i$ by independent component-wise application of the dynamics in (105) and (106) with no modification to the weights. This propagated state mixture is then updated using the method described in the following subsection.

2.5.6.1 Adaptive Update Procedure

The Kalman update is performed for each propagated state component i as follows. First, evaluate the state angle mean

$$\theta_{x_{k|k-1},i} = h_\theta(x_{k|k-1,i})$$

and variance

$$\sigma^2_{\theta,x_{k|k-1},i} = H_\theta(\hat{x}_{k|k-1,i}) P_{k|k-1,i} H_\theta(\hat{x}_{k|k-1,i})^T$$

As in the MCAEKF, use this to compute the bias significance of the measurement compared to the state using

$$C_B = \frac{r_k \sigma^2_{\theta, x_{k|k-1},i}}{2\sigma_{r,k}} \tag{123}$$

where $r_k, \sigma_{r,k}$ are the range and range standard deviation of the measurement. Furthermore, use (40) to compute the equivalent single-Gaussian KL divergence of the model. If this is less than the chosen limit of $KL^\star$, update the track component in polar space using the MCAEKF technique. If the bias significance exceeds a given threshold $C_B^\star$, then the range standard deviation of the measurement is set to $\sigma'_{r,k} = \frac{C_B}{C_B^\star}\sigma_r$ and substituted into the modified polar covariance R'_k. Then the component i is updated using

$$\hat{x}_{k|k,i} = U_{x,P}(\hat{x}_{k|k-1,i}, P_{k|k-1,i}, z_k, R'_{k,i}) \tag{124}$$

$$P_{k|k,i} = U_{P,P}(\hat{x}_{k|k-1,i}, P_{k|k-1,i}, z_k, R'_{k,i}) \tag{125}$$

If the single-Gaussian KL divergence limit is not met, then a Cartesian measurement mixture will be used to update the filter according to the conditioning process described in Section 2.4.4. With a given choice of κ, use (94) and (95) to determine the conditioning measurement mean and covariance $\theta_{y,k,i}$ and $\sigma^2_{\theta,y,k,i}$ such that the angle mean and covariance of the true measurement are updated by this conditioning measurement to $\theta_{z'_k,i} = \theta_{x_{k|k-1},i}$ and $\sigma^2_{\theta,z'_k,i} = \kappa\sigma^2_{\theta,x_{k|k-1},i}$. Then use the lookup table to generate the mixture $\left\{\omega'_{z_C,k,i,j}, z'_{C,k,i,j}, R'_{C,k,j}\right\}_j = \text{Mix}^\star_{\text{EM}}(z'_{k,i}, R'_{k,i}, KL^\star)$, where $z'_{k,i}, R'_{k,i}$ are the measurement mean and covariance conditioned in angle so the measurement is close to the support of state component i. The size of this mixture (chosen so that the KL divergence meets the limit $KL^\star$) will in general be smaller than generated by the unconditioned version of the process $\text{Mix}^\star_{\text{EM}}(z_k, R_k, KL^\star)$ due to the smaller angular variance. Update the prior component i with this mixture using

$$\hat{x}'_{k|k,i,j} = U_{x,C}(\hat{x}_{k|k-1,i}, P_{k|k-1,i}, z'_{C,k,i,j}, R'_{C,k,j}) \tag{126}$$

$$P'_{k|k,i,j} = U_{P,C}(\hat{x}_{k|k-1,i}, P_{k|k-1,i}, z'_{C,k,i,j}, R'_{C,k,j}) \tag{127}$$

and apply the inverse filtering

$$\hat{x}_{k|k,i,j} = U^{-1}_{x,\theta}(\hat{x}'_{k|k,i,j}, P'_{k|k,i,j}, \theta_{y,k,i}, \sigma^2_{\theta,y,k,i}) \tag{128}$$

$$P_{k|k,i,j} = U^{-1}_{P,\theta}(\hat{x}'_{k|k,i,j}, P'_{k|k,i,j}, \theta_{y,k,i}, \sigma^2_{\theta,y,k,i}) \tag{129}$$

to get the final posterior means and distributions of the state component i. Finally, find the weights of the posterior PDF conditioned on prior component i using

$$\tilde{\omega}_{x_{k|k},i,j} \propto \frac{\omega'_{z_C,k,i,j}}{L_{z,\theta}\left(\hat{x}_{k|k,i,j}, P_{k|k,i,j}, \theta_{y,k,i}, \sigma^2_{\theta,y,k,i}\right)} \tag{130}$$

normalized over j. See Section 2.4.4 for more details on this process.

At this point, for every prior component i, there is a posterior conditional mixture $\left\{\tilde{\omega}_{x_{k|k},i,j}, \hat{x}_{k|k,i,j}, P_{k|k,i,j}\right\}_{j=1}^{N_{G,i}}$. (If MCAEKF was used to update prior component i, then this mixture is trivially a single Gaussian with $\tilde{\omega}_{x_{k|k},i,1} = 1$). To find the final weights $\omega_{x_{k|k},i,j}$, update by the prior mixture weights and likelihoods of the measurement given the prior.

$$\omega_{x_{k|k},i,j} \propto \tilde{\omega}_{x_{k|k},i,j}\omega_{x_{k|k-1},i}L_{x,P}\left(\hat{x}_{k|k-1,i}, P_{k|k-1,i}, z_k, R_k\right) \tag{131}$$

normalized to unity over all i, j. Note that the likelihood evaluation is performed in polar space using the original unmodified polar measurements. A Cartesian mixture should not be used for this purpose as it does not represent the likelihood well at the edges of the distribution.

2.5.6.2 Mixture Component Multiplicity Control

Due to the fact that the prior state components are not updated by common measurement components in the MAGMF, the simple measurement history clustering approach used by Zhang [32] is not viable. Therefore, the MAGMF uses an adaptation of the clustering method proposed by Salmond [25].

First, relabel the indices of the mixture set $\left\{\left\{\omega_{x_{k|k},i,j}, \hat{x}_{k|k,i,j}, P_{k|k,i,j}\right\}_j\right\}_i$ to $\Theta'_{k|k} = \left\{\omega_{x_{k|k},i}, \hat{x}_{k|k,i}, P_{k|k,i}\right\}_{i=1}^{N'_{G,k|k}}$ for simplicity since the genealogy of the posterior components is not considered in this algorithm. (Note that throughout the earlier update process, the range of j is distinct for each component i.)

Second, calculate the total mean $\hat{x}_{k|k} = \mu\left(\Theta'_{k|k}\right)$ and total covariance $P_{k|k} = \Sigma\left(\Theta'_{k|k}\right)$ of the posterior PDF using (108) and (109). Calculate the equivalent polar position distribution using $\hat{z}_{k|k} = h_P(\hat{x}_{k|k})$ and $\hat{R}_{k|k} = H_P(\hat{x}_{k|k})P_{k|k}H_P(\hat{x}_{k|k})^T$. Using these polar state estimates, calculate the target number of components using the lookup table KL divergence according to

$$N^\star_{G,k|k} = N^\star_G(\hat{z}_{k|k}, \hat{R}_{k|k}, KL^\star_{pos}) \tag{132}$$

where $KL^\star_{pos}$ is a limit that may be specified independently of the $KL^\star$ used in the update.

Let the current set of mixture parameters be given by $\Theta = \{\omega_i, \mu_i, \Sigma_i\}_{i=1}^{N_G}$ and let the notation $\Theta_\Omega = \{\omega_i, \mu_i, \Sigma_i\}_{i\in\Omega}$ denote a subset of components indexed by the label set Ω. Initialize $\Theta = \Theta'_{k|k}$ from the posterior state mixture. Initialize the set of previously clustered component labels $\Omega_p = \emptyset$ and choose a cluster distance limit $D^\star$. Let $|\Theta|$ denote the cardinality of the set Θ. Then apply the following iteration until $|\Theta| \leq N^\star_{G,k|k}$.

1. If $|\Omega_p| = |\Theta_{k|k}|$ (all current components considered as principle component), reinitialize $\Omega_p = \emptyset$ and update the distance limit $D^\star \leftarrow 2D^\star$.

2. Choose the previously unselected (not in Ω_p) component index with maximum weight $i_p = \arg\max_{i\notin\Omega_p} \omega_i$ as "principle" component. Record this choice as $\Omega_p \leftarrow \Omega_p \cup \{i_p\}$.

3. Let $D(i,j) = \frac{\omega_i\omega_j}{\omega_i+\omega_j}(\mu_i - \mu_j)^T\Sigma_i^{-1}(\mu_i - \mu_j)$ be the distance cost function and

evaluate the set of components to be clustered as

$$\Omega_c = \{j \mid j \in \{1, ..., |\Theta|\} \setminus i_p, D(i_p, j) \leq D^\star\}$$

where $\setminus$ denotes the set difference operation.

4. Cluster Θ_{Ω_c} by reduction to a single Gaussian and form the updated mixture $\Theta \leftarrow \text{ReduceGM}(\Theta_{\Omega_c \cup \{i_p\}}) \cup \Theta_{\tilde{\Omega}_c \setminus \{i_p\}}$, where $\tilde{\Omega}_c$ denotes the set compliment.

5. Update the labeling of components in Ω_p by first removing all clustered indicies other than the principle and then adjusting the index labeling to account for the removed components.

When the iteration has terminated, assign $\Theta_{k|k} = \Theta$ as the final set of posterior mixture parameters of size $N^\star_{G,k|k}$ components.

2.6 Simulation of Gaussian Mixture Tracking with Contact-Lens Measurements

In order to test the new MAGMF, the two scenarios described in [28] were used. The initial conditions for the scenarios are as follows

1. Position [105 250] km, Velocity [−200 300] m/s

2. Position [1050 2500] km, Velocity [−200 300] m/s

For these scenarios, the radar measurement accuracy was $\sigma_r = 0.2$ m in range and $\sigma_\theta = 0.001$ rad in angle. This yields a bias significance value of $C_B = 0.68$ for Scenario 1, and $C_B = 6.8$ for Scenario 2. The measurement rate for the radar was 1 Hz.

For each scenario, 100 Monte Carlo runs of each filter were performed in order to collect performance statistics. For the MAGMF, a KL divergence performance limit of $KL^\star = 0.05$ was set for the measurement and $KL^\star_{pos} = 0.0125$ for the number of posterior track components. The measurement conditioning covering multiplier was

set to $\kappa = 1.5$. The CbGMF filter was allowed a maximum of 100 components before recombination of the state, and components with weight less than 0.01 times the max weight were pruned. The Zhang GMM-ITS was run with $N_G = 6$ components for Scenario 1, and $N_G = 12$ component for Scenario 2, which matches the setup in [33]. For the RPF, 100k particles were used along with a Epanechnikov kernel bandwidth of 0.5 and an effective sample threshold of 50k particles based on the results in [24].

Figures 17 through 32 show the results of the filter simulations for the two scenarios. For Scenario 1, Figure 17 shows that only the mixture and particle filters (CbGMF, GMM-ITS, MAGMF, RPF) are able to perform with range accuracy at the level of the measurements near track initialization. As expected, the converted measurement filters show a higher RMSE, especially in the biased case. However, all filters other than the biased CMF quickly settle to good performance after approximately 5 updates due to convergence of the filter in crossrange. Figure 18 shows similar performance for range rate estimation except for RPF, biased CMF, and standard EKF. For overall RMSE, Figures 19 and 20 show that the mixture filters (including MAGMF) and particle filter achieve the best estimates, except that RPF is slightly degraded in position performance and GMM-ITS is slightly degraded in velocity performance.

Due to the non-Gaussian nature of the position distribution, NEES for position is evaluated in polar space, where is is assumed the PDF more closely resembles a Gaussian. Figure 21 shows that as expected, the EKF and Biased CMF are highly optimistic. UKF and RPF show slight optimistic, and GMM-ITS is pessimistic due to the range error inflation term ($\alpha\sigma_r$ in [33]) applied to the measurement mixtures. All other mixture filters (including MAGMF), MCAEKF, and the Unbiased CMF are consistent. However, Unbiased CMF and MCAEKF trade range performance for consistency. For the (Cartesian) velocity NEES in Figure 22, all filters are slightly inconsistent at the start, likely due to the use of the one-point initialization technique

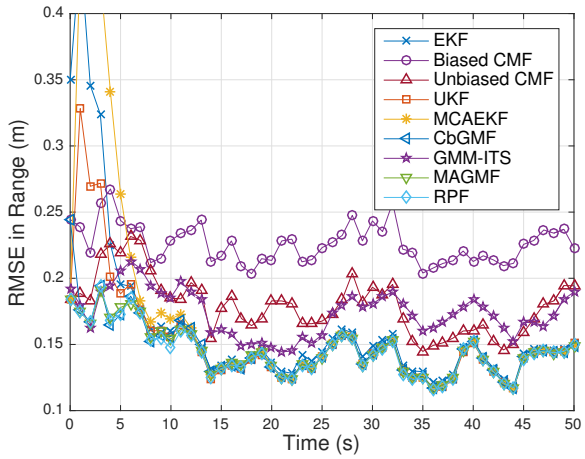

Figure 17: Range RMSE for Scenario 1

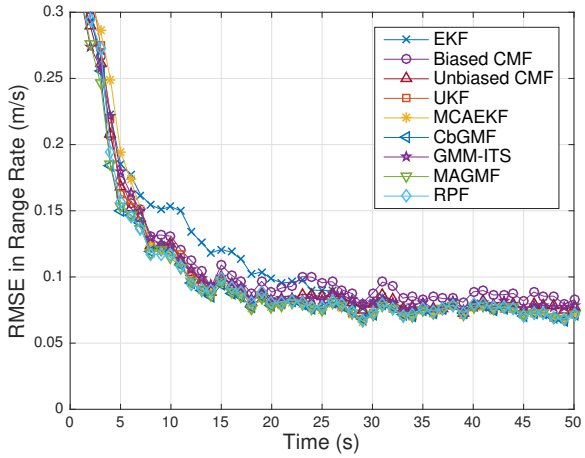

Figure 18: Range Rate RMSE for Scenario 1

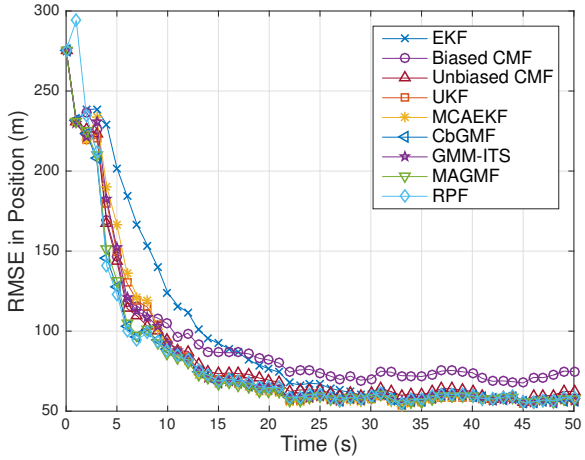

Figure 19: Position RMSE for Scenario 1

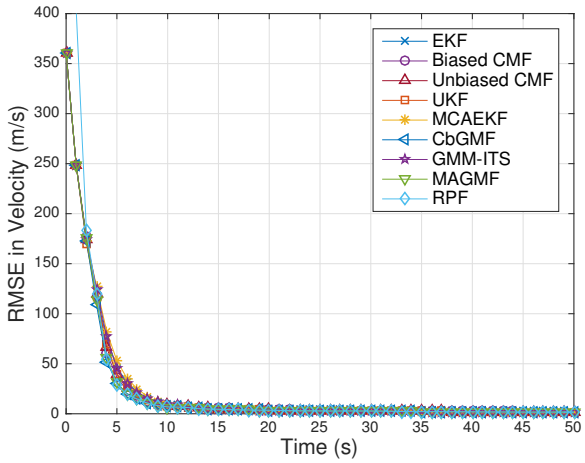

Figure 20: Velocity RMSE for Scenario 1

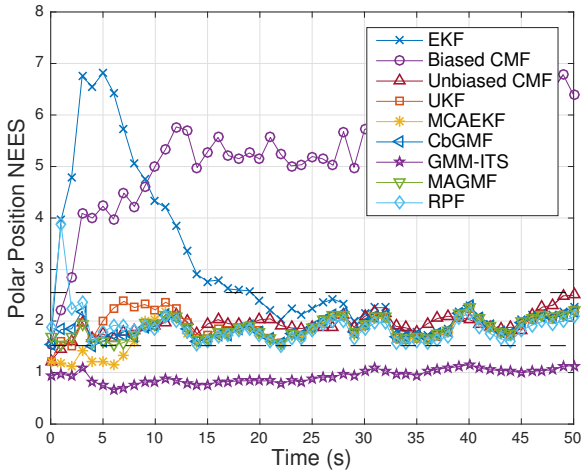

Figure 21: Mean of the Polar Position NEES for Scenario 1

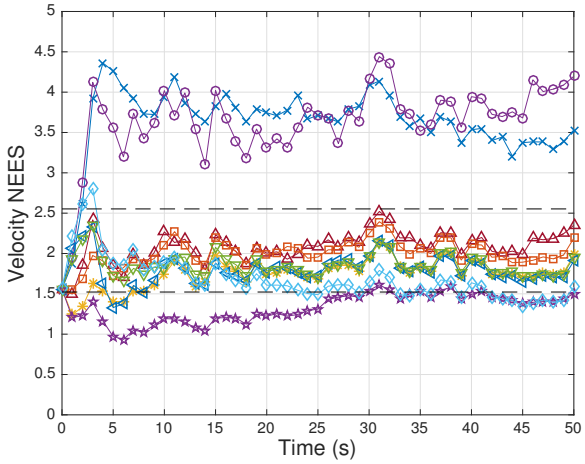

Figure 22: Mean of the Velocity NEES for Scenario 1

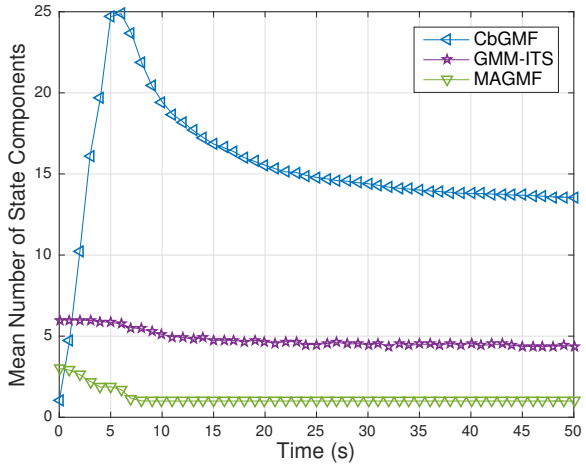

Figure 23: Mean Number of State Components for Scenario 1

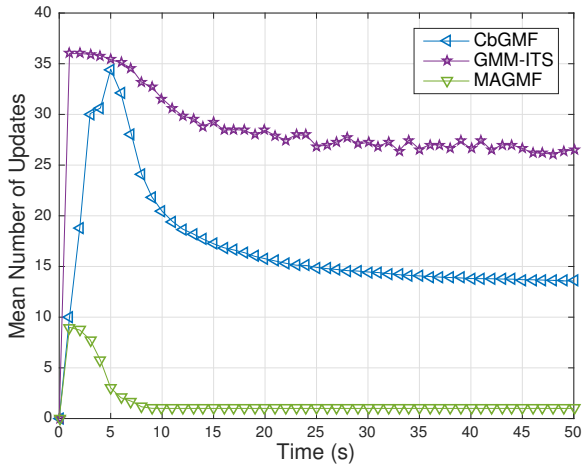

Figure 24: Mean Number of Kalman Updates for Scenario 1

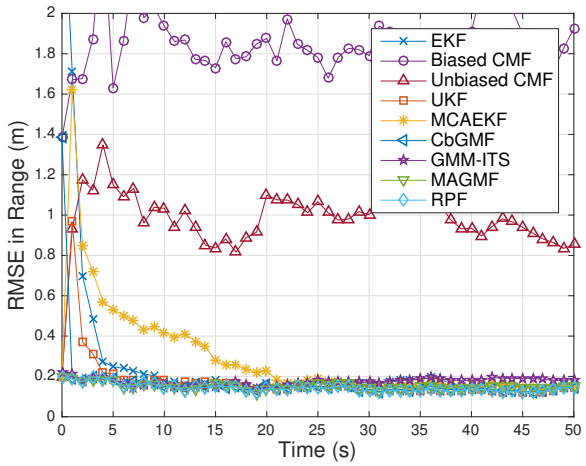

Figure 25: Range RMSE for Scenario 2

coupled with a deterministic initial velocity. Again, GMM-ITS is pessimistic and RPF, EKF, and Biased CMF are optimistic.

In order to make a comparison of computational and memory costs of the mixture filters, the mean number of posterior state components carried between iterations are shown in Figure 23, and the mean number of updates performed between measurement and state components (Cartesian product of interactions) are shown in Figure 24. (The inverse filtering updates of the MAGMF are not shown, but these occur in a scalar measurement space and therefore have an overall low computational cost). It is seen that the measurement conditioning technique used in the MAGMF along with the adaptive use of the MCAEKF inflation technique greatly lowers the number of updates that must be performed between measurement and state. The MAGMF uses 3 or fewer components to represent the state while GMM-ITS uses a peak of 6 and CbGMF over 20. At the start of the track when the bias significance is highest, it is able to provide a commensurate level of estimation performance to the GMM-ITS with half the number of state components and a fraction of the measurement-to-state updates. All single-Gaussian filters have a relative cost of 1 compared to these plots, and the RPF has a much higher cost (100k particles to maintain and resample at every step).

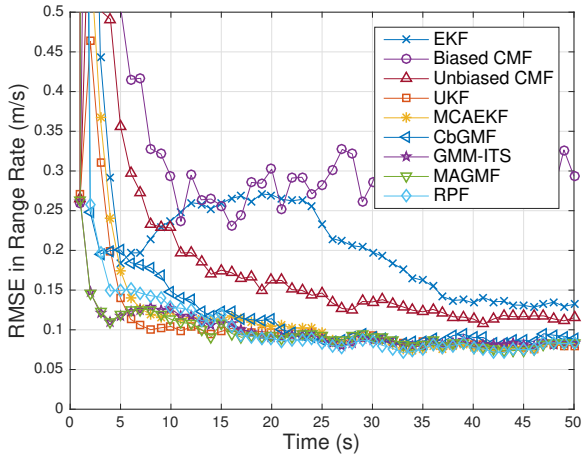

Figure 26: Range Rate RMSE for Scenario 2

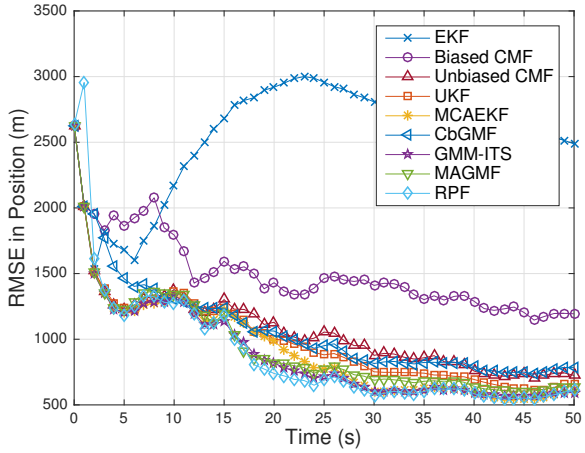

Figure 27: Position RMSE for Scenario 2

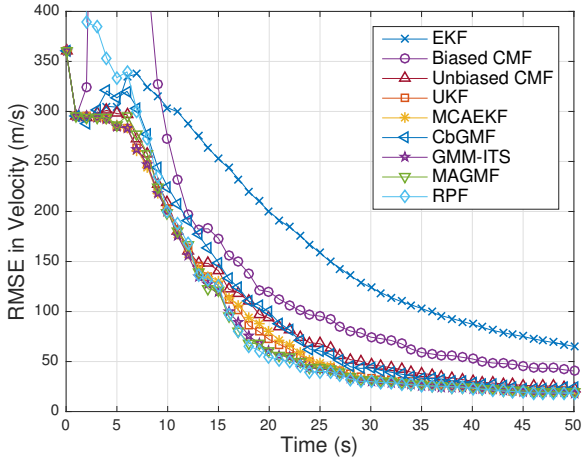

Figure 28: Velocity RMSE for Scenario 2

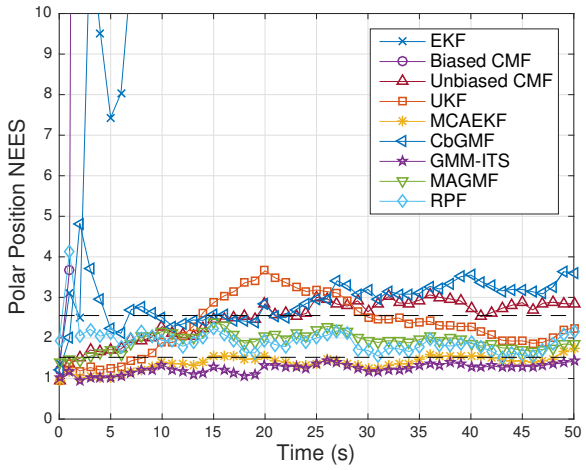

Figure 29: Mean of the Polar Position NEES for Scenario 2

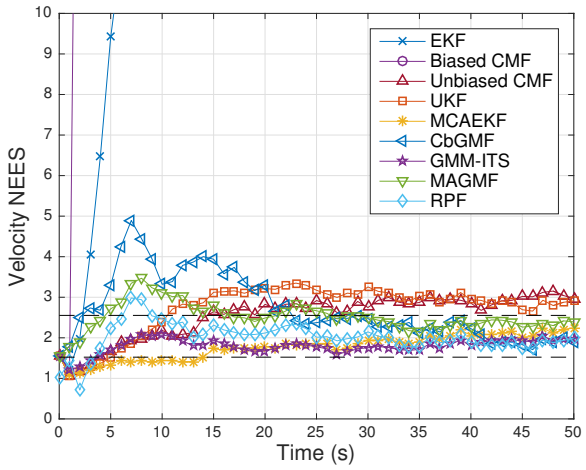

Figure 30: Mean of the Velocity NEES for Scenario 2

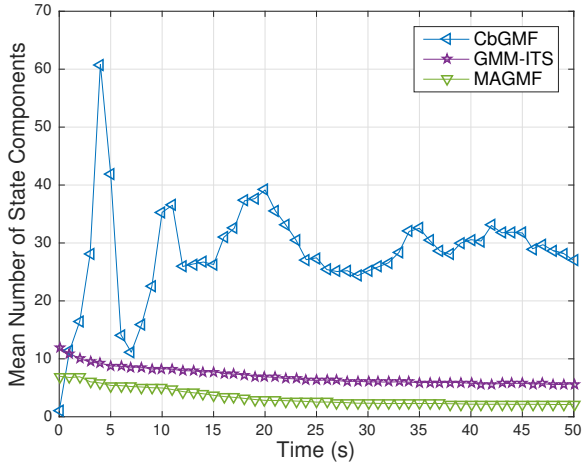

Figure 31: Mean Number of State Components for Scenario 2

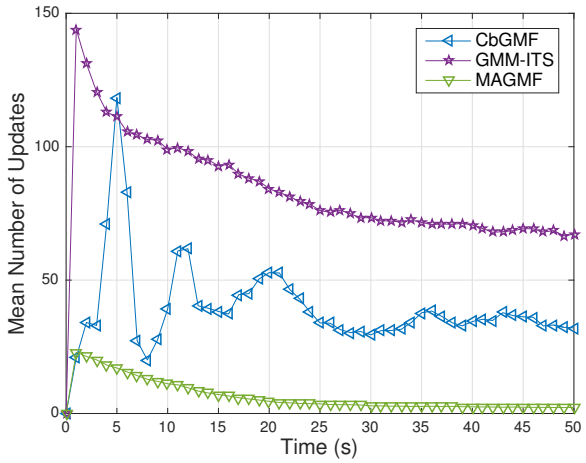

Figure 32: Mean Number of Kalman Updates for Scenario 2

For Scenario 2, Figure 25 shows that only the mixture filters and RPF are able to provide range estimation performance that is equal to or better than the measurement error right from the start of tracking. All other filters take some time to converge to this performance or never reach it at all, e.g. the CMF filters. For range rate, Figure 26 shows that only RPF, GMM-ITS, and MAGMF converge rapidly to a stable level of performance. In total position and velocity RMSE shown in Figures 27 and 28 these filters again have the best performance with RPF and MAGMF leading slightly in velocity RMSE.

For covariance consistency, Figure 29 shows MAGMF and MCAEKF with the most consistent polar position performance, while GMM-ITS is again slightly pessimistic. In velocity NEES shown in Figure 30, all filters are again slightly optimistic near track initialization, and the mixture filters and RPF converge to consistent performance over a short amount of time. In both position and velocity, EKF and Biased CMF are highly inconsistent (off the scale), and UKF is mildly optimistic.

In terms of computational complexity, Figures 31 and 32 again show that MAGMF solves the same problem as GMM-ITS with about half the state components and in this case, almost an order of magnitude fewer state-to-measurement updates due to the novel conditioning technique employed on the measurement mixture. As in

Scenario 1, the MAGMF state representation size converges steadily towards a single-Gaussian filter as the crossrange support of the state estimate converges.

2.7 Gating with Mixture Filters

One of the primary functions of a radar tracker is to associate new incoming radar measurements with existing tracks. The general procedure is to compare the state vector of existing tracks with the new measurement and evaluate the likelihood of the difference under their combined distribution. If this difference falls within a predetermined *gating region*, the measurement is considered to be an association candidate to the track. The probability that a true measurement from a given track will fall within the gating region, assuming a consistent track and measurement covariance, is known as the *probability of gating*.

A key aspect of this process is the ability to reject detections resulting from other objects under track, as well as false alarm detections. Assuming a uniform spatial probability distribution of false alarms or otherwise conflicting detections, the minimum probability of mis-association is achieved when the gating region occupies the minimum Euclidian area containing the desired probability of gating mass. Such a region is called a *tight* gating region.

Recall that a primary motivation of implementing a higher fidelity model for measurements suffering from the contact-lens problem is the effect on measurement-to-track association. Figure 2 at the beginning of this chapter showed that a moment-matched Gaussian model for these distributions is highly wasteful in terms of gating area when bias significance is high, allowing more false alarms to be gated with the track. The MAGMF takes advantage of ML GM fits of the measurement distributions to represent the filtered state more efficiently using a GM. However, it is not immediately clear how to perform the measurement gating process using a GM state representation.

2.7.1 Standard EKF Gating

In a standard EKF filter with Gaussian assumptions for measurement and track distribution, the probability of gating is defined in terms of the Mahalanobis distance of the measurement-to-track difference, which is described by (133)

$$MD = (z_k - \hat{z}_k)^T (R_k + HP_kH^T)^{-1}(z_k - \hat{z}_k) \tag{133}$$

where z_k is the measurement, $\hat{z}_k$ is the measurement estimate, and R_k and HP_kH^T are the measurement and measurement estimate covariances respectively.

For target-originated measurements, MD has a chi-squared distribution with N degrees of freedom (where N is the measurement dimension). Therefore, the boundary of a gating region with desired probability of gating can be readily determined by consulting a chi-square CDF table. Specifically $\{z_k : MD(z_k) \leq K\}$, where K is the value where the CDF of χ^2_N equals the probability of gating (P_{gate}), describes the gating region.

This gating region has two important properties.

1. It is a tight gating region

2. Its boundary is a level curve of the distribution PDF

To see the first property, consider that the Normal distribution PDF is monotonically decreasing as MD increases (MD being the argument of the exponential). Suppose the described region does not have the minimum Euclidian area. Then there must exist a set outside the region with higher probability density than in the region. However, this is false by the monotonicity of the PDF.

To see the second property, simply examine the equation of the N-d normal distribution.

$$N(x;\mu;\Sigma) = \frac{1}{\sqrt{(2\pi)^N|\Sigma|}} e^{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)} \tag{134}$$

For $x = z_k$, $\mu = \hat{z}_k$, and $\Sigma = R_k + HP_kH^T$, it is seen that the argument of the exponential is $-\frac{1}{2}MD(z_k)$. Since x appears nowhere else in the PDF, and the region boundary is a level curve of $MD(z_k)$, it is also a level curve of the PDF.

2.7.2 Gaussian Mixture Gating

In a similar way, define a gating region for an arbitrary continuous PDF. Let the gating region with probability of gating, P_{gate}, under $p(z_k)$ be defined as follows.

$$R_{gate} = \{z_k : p(z_k) \geq K\} \tag{135}$$

R_{gate} is a tight gating region, as no point outside the region has a higher density, and its boundary is a level curve of the PDF by construction. The integral $P_{gate} = \int_{R_{gate}} p(z)\mathrm{d}z$ provides a mapping between the PDF boundary value K and the desired probability of gating.

However, it is not readily apparent how to efficiently compute this integral for an arbitrary GM. In [8], expensive numerical sampling was required to achieve reasonable results for gating in Cartesian space. This method is not desirable for use in an online filter.

An useful approximation is made by noting that the position subspace of the state vector should be nearly Gaussian in the original measurement space of range and angle due to the geometry of the measurements. Since the measurement itself is Gaussian in this space, the gate Mahalanobis distance may be computed in polar coordinates (as in a standard EKF). Let $\left\{\omega_{x,k|k-1,i}, \hat{x}_{k|k-1,i}, P_{k|k-1,i}\right\}_i$ denote the prior state mixture representation at time k. Then using the measurement function h_P along with the Unscented Transform, convert each component of this mixture to polar space.

$$\left(\hat{z}_{k|k-1,i}, P_{\hat{z},k|k-1,i}\right) = \mathrm{UT}\left(h_P, \hat{x}_{k|k-1,i}, P_{k|k-1,i}\right) \tag{136}$$

Form a single-Gaussian approximation of this polar mixture using moment matching.

$$\left(\hat{z}_{k|k-1}, P_{\hat{z},k|k-1}\right) = \mathrm{ReduceGM}\left(\left\{\omega_{x,k|k-1,i}, \hat{z}_{k|k-1,i}, P_{\hat{z},k|k-1,i}\right\}_i\right) \tag{137}$$

Finally, compute the Mahalanobis distance in polar space.

$$MD_k = (z_k - \hat{z}_{k|k-1})^T (R_k + P_{\hat{z},k|k-1})^{-1} (z_k - \hat{z}_{k|k-1}) \tag{138}$$

This distance is then compared to Chi-squared tables for gating as in the standard EKF.

2.7.3 Simulation of Gating Approach

In order to demonstrate this approach, the gating process is applied to the measurements used in the MAGMF filter simulations in Section 2.6. For simplicity, the gate will be evaluated at time k assuming perfect association decisions before that time. Therefore, the state prior used in the gate will incorporate all previous measurements. All filters with a Gaussian mixture state representation are evaluated according to the method described in the previous subsection, while all single-Gaussian filters use the standard EKF gating approach.

The size of the gating region is evaluated in terms of the square root of the joint covariance matrix determinant, as this is proportional to the geometric area inside the ellipse. Assuming a constant false alarm density, this area is proportional to the expected number of false alarms that will be gated. For convenience, the gating area is normalized relative to the area in the measurement covariance ellipse alone. This normalized area is given by

$$\hat{A}_k = \sqrt{\frac{\left|R_k + P_{\hat{z},k|k-1}\right|}{|R_k|}} \tag{139}$$

The other key metric association with the gating process is the probability that a true measurement will gate. This is evaluated in terms of the fraction of Monte Carlo trials out of the 100 performed for which the measurement Mahanalobis distance MD_k falls within the Chi-squared score of choice. Here a value of $K = 5.991$ is used, which represent a 95% cumulative distribution function (CDF) score for a Chi-squared distribution with 2 degrees of freedom.

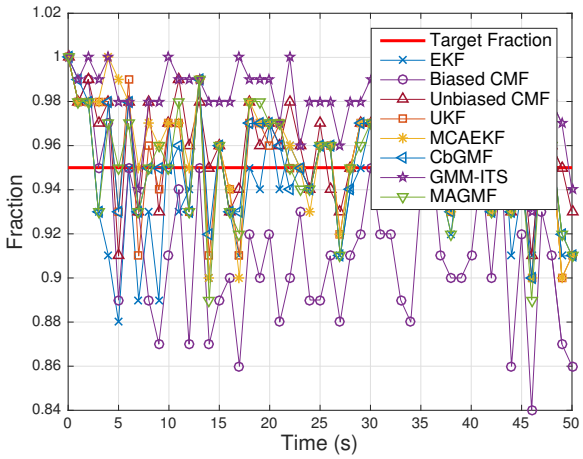

Figure 33: Fraction of Measurements in Gate for Scenario 1

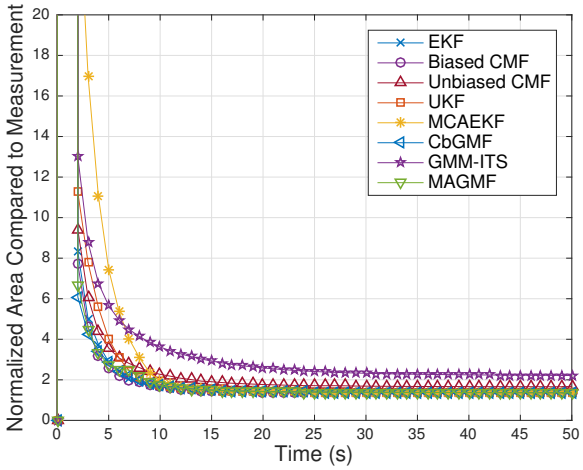

Figure 34: Gating Region Size for Scenario 1

Figure 33 shows the gating performance of the true measurement for Scenario 1 with $C_B = 0.6$. All filters are able to gate the true measurement with close to 95% probability. However, the biased CMF has slightly degraded performance relative to the other filters, as does the EKF near the beginning of the track. This is expected, as it is exactly the problem that the debiased transform is designed to solve for measurements with moderate bias significance.

To quantify false alarm performance, Figure 34 shows the normalized gate area ($\hat{A}$) of each filter. The EKF, Biased CMF , CbGMF, and MAGMF all have a very similar gate size which is less that the other filters. The UKF and Unbiased CMF have

a slightly larger gate area early in the track, and the MCAEKF has a significantly larger gate area. However, these filters converge to a small gate area over time. The GMM-ITS uses a range variance inflation in its GM modeling approach and therefore has an increased gate area for all time.

Combining the results of these two figures, the CbGMF and MAGMF are the only two filters which match the desired true target gating probability while having the smallest gating area early in the track. Based on the results of Section 2.6, the MAGMF is able to do this with a lower number of mixture components.

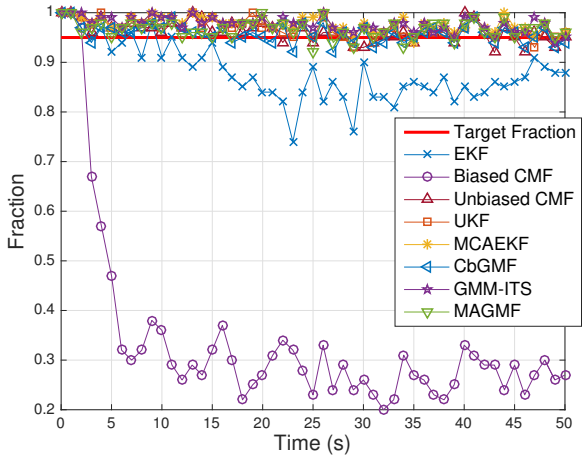

Figure 35: Fraction of Measurements in Gate for Scenario 2

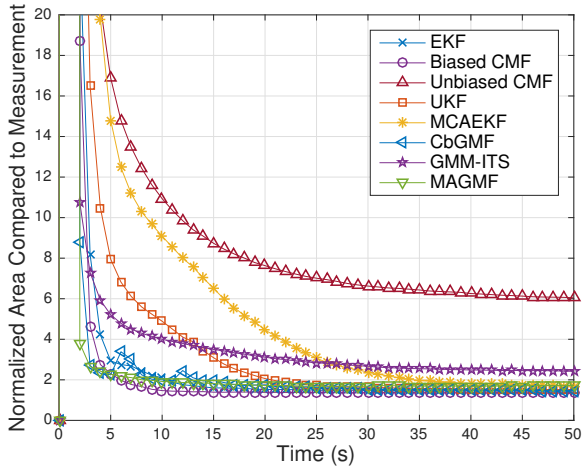

Figure 36: Gating Region Size for Scenario 2

This difference in performance is even more pronounced for Scenario 2 with $C_B =$

6. Figure 35 shows the fraction of MC trials for which the true measurement is gated given a target fraction of 95%. The Biased CMF is highly inconsistent and is able to gate the true measurement less than half the time once the filter converges. The EKF is also inconsistent and suffers a loss of 5-10% in gating probability. All other filters are able to gate the true measurement close to the desired probability of 95%.

Figure 36 shows the relative area in the gating ellipses associated with these filters for Scenario 2. The differences in these areas are much more pronounced than in Scenario 1. The MAGMF with the best performance is followed by the CbGMF, Biased CMF, and EKF. However, as seen in Figure 35, the Biased CMF and EKF fail to gate the true measurement reliably. All other filters have a significantly large gate area size near the start of the track, and the GMM-ITS has an inflated gate size throughout the track. Only the MAGMF and CbGMF are able to gate the true measurements and provide an efficient gating region in terms of area. By referencing Figures 31 and 32, it is seen that the MAGMF performs this task with far fewer components in the state representation and fewer measurement-to-state updates in the filter.

2.8 Conclusions

In this chapter, an information-theoretic interpretation of the contact-lens problem was developed which provides the ability to link existing metrics of performance in the literature to common statistical techniques such as ML. A technique for generating an ML Gaussian mixture representation of a polar measurement distribution in Cartesian space was presented and shown to be effective based on numerical sampling simulation. The ML parameters are stored in a compact lookup table for efficient use. Furthermore a novel conditioning process applied to the measurement before conversion to Cartesian space was shown to reduce the size of the mixture necessary to model the measurement and improve overall estimation performance.

This ML modeling approach and preconditioning process were then combined with existing GM filtering techniques in the new MAGMF filter. Figure 37 summarizes the performance of each filter studied in tracking the most difficult scenario. Comparison of the MAGMF to alternative approaches shows that it has excellent range estimation performance while preserving covariance consistency. Furthermore, it is able to achieve this performance at a relatively low computational cost compared to particle filters and existing GM filters proposed as solutions to the contact-lens problem.

Finally, the MAGMF was compared to these other filters in terms of measurement-to-track gating performance. Again, Figure 37 shows the gating results of the MAGMF relative to the other filters studied. The results show that the technique of single-Gaussian moment matching to a predicted measurement GM in the polar measurement space is effective in gating true measurements when the high-quality mixture estimates from the MAGMF are used. Additionally, the gating region size of the MAGMF is significantly smaller compared to single-Gaussian approaches when the bias significance is high, and this reduces the probability of mis-associating a false alarm to the track. Much of the contents of this chapter are represented in a submitted journal paper [12] and two conference papers [8] and [9].

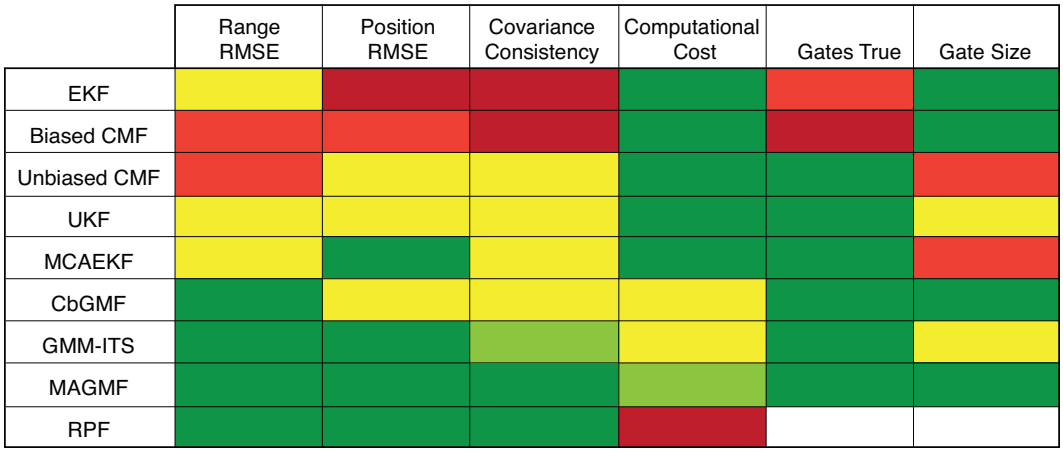

Figure 37: Performance Summary of Filters in 2D Monostatic Simulations

CHAPTER III

GM FILTERING FOR 2D BISTATIC RADARS

A bistatic radar consists of a transmitter (Tx) and receiver (Rx) separated by a distance known as the "baseline" (L). Figure 38 shows such a radar system laid out according to the geometric conventions described in [31]. R_R denotes the range from the receiver to the target, R_T denotes the range from the transmitter to the target, and θ_R denotes the angle of arrival of the target at the receiver relative to the y-axis. The ellipse pictured is an iso-bistatic-range contour of the target position, i.e., $R_S = R_T + R_R$ is constant for all points on the ellipse. Note that the transmitter and receiver are located at the foci of the ellipse.

Like monostatic radar, bistatic radar systems experience distortion of their measurement uncertainty due to the non-linear transformation between measurement space and state space. However, a metric to quantify when this non-linearity becomes severe similar to the bias significance metric for monostatic radar does not exist in the literature. Most previous attempts to address problems with bistatic radar measurements have focused on single-Gaussian solutions. A set of debiased transform equations was developed by Coogle [4] which provides an accurate moment

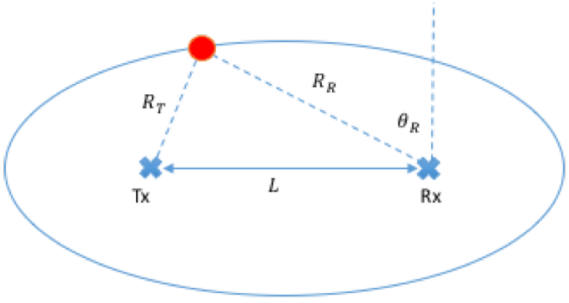

Figure 38: 2D Bistatic Radar Geometry

match to converted bistatic radar measurements as [20] did for monostatic radar. Both Coogle [5] and Crouse [6] studied the use of UKF and sigma point filters applied to bistatic radar. No examples of the use of Gaussian mixtures for tracking with a bistatic radar were found in the literature.

The objective of this chapter is to extend the methods applied to 2D monostatic radar in Chapter 2 to a 2D bistatic radar. Since no equivalent bias significance metric exists for bistatic radar measurements, the first step is to derive a similar metric which can be used to drive further analysis in the context of EM mixture fitting. After this is established, ML GM fits to the bistatic measurement are studied with the goal of generating an efficient lookup table for the ML GM parameters of component weight, mean and covariance. However, the 2D monostatic case is aided by its circular geometry which allows straightforward removal of one degree of freedom (mean angle). This is not possible for the elliptical geometry of the bistatic radar. A significant portion of the research presented in this chapter can be found in [10].

3.1 Study of the Transform

The natural measurement space for a bistatic radar is bistatic range sum $R_S = R_T + R_R$ and angle of arrival from the receiver θ_R. These measurements are converted to a Cartesian coordinate system centered on the midpoint of the transmitter-to-receiver baseline using [31]

$$x = \frac{L}{2} + R_R \sin\theta_R \qquad\qquad y = R_R \cos\theta_R \tag{140}$$

$$R_R = \frac{R_S^2 - L^2}{2\left(R_S + L\sin\theta_R\right)} = R_S \frac{1 - e^2}{2\left(1 + e\sin\theta_R\right)} \tag{141}$$

where $e = \frac{L}{R_S}$ is the eccentricity of the iso-range bistatic ellipse.

For convenience, define $x_C = \begin{bmatrix} x \\ y \end{bmatrix}$ and $x_B = \begin{bmatrix} R_S \\ \theta_R \end{bmatrix}$. Then write this transform as a

vector function according to

$$x_C = \text{B2C}(x_B) = \begin{bmatrix} \frac{L}{2} + R_R \sin\theta_R \\ R_R \cos\theta_R \end{bmatrix} \tag{142}$$

The Jacobian is

$$\frac{\partial}{\partial x_B}\text{B2C} = J_{\text{B2C}}(x_B) = \begin{bmatrix} \frac{\partial R_R}{\partial R_S}\sin\theta_R & \frac{\partial R_R}{\partial \theta_R}\sin\theta_R + R_R\cos\theta_R \\ \frac{\partial R_R}{\partial R_S}\cos\theta_R & \frac{\partial R_R}{\partial \theta_R}\cos\theta_R - R_R\sin\theta_R \end{bmatrix} \tag{143}$$

where $\frac{\partial R_R}{\partial R_S}$ and $\frac{\partial R_R}{\partial \theta_R}$ are shown in full in the subsequent section. The Jacobian determinant is

$$|J_{\text{B2C}}| = R_S \frac{\partial R_R}{\partial R_S} \tag{144}$$

which is important in evaluating the KL divergence of PDF approximations in Cartesian space (see Section 2.2.2).

The inverse transformation is

$$x_B = \text{C2B}(x_C) = \begin{bmatrix} R_R + R_T \\ \tan^{-1}\frac{x-\frac{L}{2}}{y} \end{bmatrix} \tag{145}$$

with Jacobian

$$\frac{\partial}{\partial x_B}\text{C2B} = J_{\text{C2B}}(x_C) = \begin{bmatrix} \frac{x-\frac{L}{2}}{R_R} + \frac{x+\frac{L}{2}}{R_T} & \frac{y}{R_R} + \frac{y}{R_T} \\ \frac{y}{R_R^2} & -\frac{x-\frac{L}{2}}{R_R^2} \end{bmatrix} \tag{146}$$

where

$$R_R = \sqrt{\left(x - \frac{L}{2}\right)^2 + y^2} \qquad R_T = \sqrt{\left(x + \frac{L}{2}\right)^2 + y^2} \tag{147}$$

are the range to receiver and range to transmitter expressed in Cartesian coordinates.

3.1.1 Derivatives of the Transform

From [31] and further manipulations, the first derivatives of the bistatic range transformation are

$$\frac{\partial R_R}{\partial R_S} = \frac{1 + e^2 + 2e\sin\theta_R}{2\left(1 + e\sin\theta_R\right)^2} \tag{148}$$

$$\frac{\partial R_R}{\partial \theta_R} = R_S \frac{e\cos\theta_R(1-e^2)}{2\left(1+e\sin\theta_R\right)^2} = R_R \frac{e\cos\theta_R}{(1+e\sin\theta_R)} \tag{149}$$

Furthermore, the second derivatives are

$$\frac{\partial^2 R_R}{\partial {R_S}^2} = -\frac{1}{R_S}\frac{e^2\cos^2\theta_R}{2\left(1+e\sin\theta_R\right)^3} \tag{150}$$

$$\begin{aligned}\frac{\partial^2 R_R}{\partial {\theta_R}^2} &= \frac{1}{2}R_S e(1-e^2)\frac{\sin\theta_R + e\left(1+\cos^2\theta_R\right)}{2\left(1+e\sin\theta_R\right)^3} \\ &= R_R\frac{e\sin\theta_R + e^2\left(1+\cos^2\theta_R\right)}{2\left(1+e\sin\theta_R\right)^3}\end{aligned} \tag{151}$$

$$\frac{\partial^2 R_R}{\partial R_S \partial \theta_R} = -e^2\cos\theta_R\frac{\sin\theta_R + e}{\left(1+e\sin\theta_R\right)^3} \tag{152}$$

These expressions for the derivatives are convenient, as they reveal dependencies on the overall bistatic range of the problem (R_S) rather than normalized parameters such as the eccentricity e. In particular, the first and second derivatives of receiver range with respect to angle are scaled directly by the receiver range (which has the same order as the range sum). Therefore, normalized versions of these derivatives can be defined as follows.

$$\frac{\partial \widehat{R}_R}{\partial \theta_R} = \frac{1}{R_R}\frac{\partial R_R}{\partial \theta_R} \qquad\qquad \frac{\partial^2 \widehat{R}_R}{\partial {\theta_R}^2} = \frac{1}{R_R}\frac{\partial^2 R_R}{\partial {\theta_R}^2} \tag{153}$$

3.1.2 Statistics of the Converted Bistatic Radar Measurement

Let the measurement errors be Gaussian with $R_S \sim N\left(\bar{R}_S, \sigma_S^2\right)$ and $\theta_R \sim N\left(\bar{\theta}_R, \sigma_\theta^2\right)$. The nonlinear transformation of these Gaussian errors using Equations (140) and (141) results in a non-Gaussian Cartesian error distribution. However, this error is frequently approximated by a single Gaussian distribution with moments matched to the true distribution. Here the mean and covariance of the Cartesian distribution are computed using first and second order Taylor expansion of the transformation in a similar way to Chapter 2 for monostatic radar and [27] for monostatic RUV coordinates.

In order to remove the effect of the mean receiver angle from the equations, perform the following coordinate transform.

$$U = \begin{bmatrix} \sin\bar{\theta}_R & \cos\bar{\theta}_R \\ -\cos\bar{\theta}_R & \sin\bar{\theta}_R \end{bmatrix} \qquad \begin{bmatrix} x' \\ y' \end{bmatrix} = U \begin{bmatrix} x \\ y \end{bmatrix} \tag{154}$$

leading to

$$x' = R_R \cos\tilde{\theta}_R + \frac{L}{2}\sin\bar{\theta}_R \qquad y' = -R_R \sin\tilde{\theta}_R - \frac{L}{2}\cos\bar{\theta}_R \tag{155}$$

where $\tilde{\theta}_R = \theta_R - \bar{\theta}_R$ is the zero-mean error in angle.

Now, in an analogous way to [4] and [27], approximations for the mean and covariance of the converted measurement are calculated based on Taylor expansion of the conversion equations. For brevity, receiver range and associated derivatives are assumed to be evaluated at the distribution mean, and the explicit notation for the evaluation point (*i.e.,* $R_R = R_R|_{\bar{R}_S,\bar{\theta}_R}$, $\frac{\partial R_R}{\partial \theta_R} = \left.\frac{\partial R_R}{\partial \theta_R}\right|_{\bar{R}_S,\bar{\theta}_R}$, *etc*) is suppressed.

3.1.3 First Order Mean and Covariance Approximations

A first order Taylor expansion of the rotated coordinates is

$$\begin{aligned} x' &\approx R_R + \left(\frac{\partial R_R}{\partial \theta_R}\right)\tilde{\theta}_R + \left(\frac{\partial R_R}{\partial R_S}\right)\tilde{R}_S + \frac{L}{2}\sin\bar{\theta}_R \\ y' &\approx -R_R\tilde{\theta}_R - \frac{L}{2}\cos\bar{\theta}_R \end{aligned} \tag{156}$$

This expansion leads to the following approximations of mean and covariance.

$$E\{x'\} \approx \mu_{x',1} = R_R + \frac{L}{2}\sin\bar{\theta}_R \qquad E\{y'\} \approx \mu_{y',1} = -\frac{L}{2}\cos\bar{\theta}_R \tag{157}$$

with $\mu_1 = U^T \begin{bmatrix} \mu_{x',1} & \mu_{y',1} \end{bmatrix}^T$.

$$\begin{aligned} Var\{x'\} &\approx C_{x',x',1} = \left(\frac{\partial R_R}{\partial R_S}\right)^2 \sigma_S^2 + \left(\frac{\partial R_R}{\partial \theta_R}\right)^2 \sigma_\theta^2 \\ Var\{y'\} &\approx C_{y',y',1} = R_R^2\sigma_\theta^2 \\ Cov\{x',y'\} &\approx C_{x',y',1} = R_R\left(\frac{\partial R_R}{\partial R_S}\right)\sigma_\theta^2 \end{aligned} \tag{158}$$

The full covariance matrix approximation is

$$\Sigma_1 = U^T \begin{bmatrix} C_{x',x',1} & C_{x',y',1} \\ C_{x',y',1} & C_{y',y',1} \end{bmatrix} U \tag{159}$$

Since the determinant of the true covariance matrix is given by

$$|\Sigma| = Var\{x'\} Var\{y'\} - Cov\{x', y'\}^2 \tag{160}$$

and $|U| = 1$, the true covariance determinant is approximated by

$$|\Sigma| \approx |\Sigma_1| = R_R^2 \left(\frac{\partial R_R}{\partial R_S}\right)^2 \sigma_S^2 \sigma_\theta^2 \tag{161}$$

Let CB1 denote the model where the measurement distribution is represented by a single Gaussian with mean μ_1 and covariance Σ_1 (in analogy to CM1 from [27]).

3.1.4 Second Order Mean and Covariance Approximations

In order to achieve a higher-order approximation, it is convenient to Taylor expand the receiver range to second order while leaving the sin and cos terms intact, as exact moments of these functions are available in [19].

$$\begin{aligned} R_R \approx R_R &+ \left(\frac{\partial R_R}{\partial R_S}\right) \tilde{R}_S + \left(\frac{\partial R_R}{\partial \theta_R}\right) \tilde{\theta}_R \\ &+ \left(\frac{\partial^2 R_R}{\partial {R_S}^2}\right) \frac{\tilde{R}_S^2}{2} + \left(\frac{\partial^2 R_R}{\partial {\theta_R}^2}\right) \frac{\tilde{\theta}_R^2}{2} + \left(\frac{\partial^2 R_R}{\partial R_S \partial \theta_R}\right) \tilde{R}_S \tilde{\theta}_R \end{aligned} \tag{162}$$

Based on this expansion combined with (140) and (141), the means are approximated by

$$\begin{aligned} E\{x'\} \approx \mu_{x',2} &= \left[R_R + \left(\frac{\partial^2 R_R}{\partial {R_S}^2}\right) \frac{\sigma_S^2}{2} + \left(\frac{\partial^2 R_R}{\partial {\theta_R}^2}\right) \frac{\sigma_\theta^2}{2}\right] e^{-\frac{\sigma_\theta^2}{2}} + \frac{L}{2} \sin \bar{\theta}_R \\ E\{y'\} \approx \mu_{y',2} &= -\left(\frac{\partial R_R}{\partial \theta_R}\right) \sigma_\theta^2 e^{-\frac{\sigma_\theta^2}{2}} - \frac{L}{2} \cos \bar{\theta}_R \end{aligned} \tag{163}$$

and $\mu_2 = U^T \begin{bmatrix} \mu_{x',2} & \mu_{y',2} \end{bmatrix}^T$. After a large amount of algebra which is omitted for brevity, it can be shown that

$$Var\{x'\} \approx C_{x',x',2} = \left(\frac{\partial R_R}{\partial R_S}\right)^2 \sigma_S^2 + \left(\frac{\partial R_R}{\partial \theta_R}\right)^2 \sigma_\theta^2 + X_4 R_R^2 \sigma_\theta^4$$

$$Var\{y'\} \approx C_{y',y',2} = R_R^2 \sigma_\theta^2 + Y_4 R_R^2 \sigma_\theta^4$$

$$Cov\{x', y'\} \approx C_{x',y',2} = R_R \left(\frac{\partial R_R}{\partial R_S}\right) \sigma_\theta^2 + C_4 R_R^2 \sigma_\theta^4$$

where

$$X_4 = \frac{1}{2} - 3\left(\frac{\partial \widehat{R}_R}{\partial \theta_R}\right)^2 - \left(\frac{\partial^2 \widehat{R}_R}{\partial {\theta_R}^2}\right) + \frac{1}{2}\left(\frac{\partial^2 \widehat{R}_R}{\partial {\theta_R}^2}\right)^2$$

$$Y_4 = 3\left(\frac{\partial^2 \widehat{R}_R}{\partial {\theta_R}^2}\right) + 2\left(\frac{\partial \widehat{R}_R}{\partial \theta_R}\right)^2 - 1$$

$$C_4 = -3\left(\frac{\partial \widehat{R}_R}{\partial \theta_R}\right) + \frac{5}{2}\left(\frac{\partial \widehat{R}_R}{\partial \theta_R}\right)\left(\frac{\partial^2 \widehat{R}_R}{\partial {\theta_R}^2}\right)$$

using the normalized version of the angle partials from (153).

The full covariance matrix approximation is

$$\Sigma_2 = U^T \begin{bmatrix} C_{x',x',2} & C_{x',y',2} \\ C_{x',y',2} & C_{y',y',2} \end{bmatrix} U \tag{164}$$

Let CB2 denote the model where the measurement distribution is represented by a single Gaussian with mean μ_2 and covariance Σ_2.

With further algebra, the determinant of the covariance matrix is approximated as

$$|\Sigma| \approx |\Sigma_2| = R_R^2 \left(\frac{\partial R_R}{\partial R_S}\right)^2 \sigma_S^2 \sigma_\theta^2 + K_4 R_R^4 \sigma_\theta^6 = |\Sigma_1| \left(1 + \frac{K_4 R_R^2 \sigma_\theta^4}{\left(\frac{\partial R_R}{\partial R_S}\right)^2 \sigma_S^2}\right) \tag{165}$$

where

$$\begin{aligned} K_4 &= X_4 - 2\left(\frac{\partial \widehat{R}_R}{\partial \theta_R}\right) C_4 + \left(\frac{\partial \widehat{R}_R}{\partial \theta_R}\right)^2 Y_4 \\ &= \frac{1}{2}\left[1 - \left(\frac{\partial^2 \widehat{R}_R}{\partial {\theta_R}^2}\right)\right]^2 + 2\left(\frac{\partial \widehat{R}_R}{\partial \theta_R}\right)^2 \left[\left(\frac{\partial \widehat{R}_R}{\partial \theta_R}\right)^2 - \left(\frac{\partial^2 \widehat{R}_R}{\partial {\theta_R}^2}\right) + 1\right] \end{aligned} \tag{166}$$

Further defining $R_{RS} = \frac{R_R}{R_S}$ gives

$$\frac{\partial \widehat{R}_R}{\partial R_\theta} = -\sqrt{\frac{e^2 - (2R_{RS} - 1)^2}{1 - e^2}} \qquad \frac{\partial^2 \widehat{R}_R}{\partial {R_\theta}^2} = -1 + \frac{2R_{RS}\,(3 - 4R_{RS})}{1 - e^2} \tag{167}$$

which yields

$$K_4 = \frac{2R_{RS}}{(1 - e^2)^2} \tag{168}$$

Additionally,

$$\frac{\partial R_R}{\partial R_S} = 2R_{RS}\frac{1 - R_{RS}}{1 - e^2} \tag{169}$$

so

$$|\Sigma_2| = |\Sigma_1| \left(1 + \left(\frac{R_{RS}}{1 - R_{RS}}\right)^2 \frac{R_S^2 \sigma_\theta^4}{2\sigma_S^2}\right) \tag{170}$$

3.1.5 Relationship to Condition Ratio

In [27], a comparison of the CM1 (first order Taylor) and CM2 (second order Taylor) conversions of RUV measurements to Cartesian coordinates is made in terms of the ratio of condition numbers (condition ratio) of the covariance matrices. The condition number $\kappa(M)$ of a matrix M is defined as the ratio of the largest to the smallest eigenvalue. The condition number ratio indicator from [27] is given by

$$\frac{\kappa(R_2)}{\kappa(R_1)} \approx \left(1 + \frac{3}{2} r_m^2 \frac{\sigma_u^4}{\sigma_r^2} \frac{3 - 3u_1^2 + u_1^4}{1 - 2u_1^2 + u_1^4}\right)^{-1} \tag{171}$$

where $u_1 = \sqrt{u_m^2 + v_m^2}$. R_1 and R_2 are the first and second order Cartesian covariance approximations, r_m, u_m, and v_m are the means of the range, U, and V measurements, and σ_r, σ_u, and $\sigma_v = \sigma_u$ are the error standard deviations of the measurements.

For the 2D bistatic case, only two dimensions and two eigenvalues are involved. If the largest eigenvalue is assumed approximately constant between the two approximations and the most significant change is in the smallest eigenvalue, then the condition ratio is approximated by the ratio of the determinants.

$$\frac{\kappa(\Sigma_2)}{\kappa(\Sigma_1)} \approx \frac{|\Sigma_1|}{|\Sigma_2|} = \left(1 + \left(\frac{R_{RS}}{1 - R_{RS}}\right)^2 \frac{R_S^2 \sigma_\theta^4}{2\sigma_S^2}\right)^{-1} \tag{172}$$

which has a similar form to (171). As in [27], this condition ratio can also be used to evaluate whether CB1 or CB2 should be used for a given bistatic measurement. If the condition ratio is too low, the CB2 conversion should be used rather than CB1.

3.2 Bias Significance and KL Divergence applied to Bistatic Radar

Given a mean range of $\bar{r}$, a range error of σ_r and and angle error of σ_θ, the key performance indicator of bias significance for a monostatic radar is given by (1). Monostatic radar is a special case of bistatic radar, so the preceding results for approximation of the bistatic distribution moments can be reduced to a monostatic case by substituting $R_S = 2r$, $\theta_R = \theta$, $R_R = r$. Additionally, $\frac{\partial R_R}{\partial R_S} = R_{RS} = 1/2$ and $\frac{\partial R_R}{\partial R_\theta} = 0$. Therefore,

$$|\Sigma_1| = \bar{r}^2 \sigma_r^2 \sigma_\theta^2 \tag{173}$$

$$|\Sigma_2| = |\Sigma_1| \left(1 + \left(\frac{\bar{r}^2 \sigma_\theta^4}{2\sigma_r^2}\right)\right) = |\Sigma_1| \left(1 + 2C_B^2\right) \tag{174}$$

With this observation that the bias significance can be directly related to the inflation in covariance (or equivalently, the condition ratio) between CB1 and CB2 in the monostatic case, the following extension is made to the definition of bias significance so that it applies also to the bistatic case.

$$C_B = \left(\frac{R_{RS}}{1 - R_{RS}}\right) \frac{R_S \sigma_\theta^2}{2\sigma_S} \tag{175}$$

3.2.1 KL Divergence of Bistatic Measurement

For the bistatic measurement transformation of (140), the Jacobian is computed as follows

$$\begin{aligned} \left|\frac{\partial f}{\partial x}\right| &= \begin{vmatrix} \frac{\partial R_R}{\partial R_S} \sin\theta_R & R_R \cos\theta + \frac{\partial R_R}{\partial \theta_R} \sin\theta_R \\ \frac{\partial R_R}{\partial R_S} \cos\theta_R & -R_R \sin\theta + \frac{\partial R_R}{\partial \theta_R} \cos\theta_R \end{vmatrix} \\ &= R_R \frac{\partial R_R}{\partial R_S} \end{aligned} \tag{176}$$

where here x denotes the vector $x = \begin{bmatrix} x & y \end{bmatrix}^T$.

Let the model $q(x)$ be given by CB2. Then, the cross-entropy is

$$\begin{aligned} H(p,q) = \ln(2\pi) + \frac{1}{2}\ln|\Sigma_2| - E_p\left\{\ln\left(R_R\frac{\partial R_R}{\partial R_S}\right)\right\} \\ + \frac{1}{2}E_p\left\{(x-\mu_2)^T{\Sigma_2}^{-1}(x-\mu_2)\right\} \end{aligned} \tag{177}$$

By the permutation property of the trace,

$$(x-\mu_2)^T{\Sigma_2}^{-1}(x-\mu_2) = \mathrm{tr}\left((x-\mu_2)(x-\mu_2)^T{\Sigma_2}^{-1}\right) \tag{178}$$

If μ_2 and Σ_2 are very close to the true mean and covariance (the second order approximation is a good one), then $E_p\left\{(x-\mu_2)(x-\mu_2)^T\right\} \approx \Sigma \approx \Sigma_2$ so that $E_p\left\{\mathrm{tr}\left((x-\mu_2)(x-\mu_2)^T{\Sigma_2}^{-1}\right)\right\} \approx \mathrm{tr}(I_2) = 2$.

The entropy of the original bistatic Gaussian measurement is

$$H(p) = 1 + \ln(2\pi) + \ln(\sigma_S\sigma_\theta) \tag{179}$$

Combining (176) and (179) yields

$$\begin{aligned} D_{KL}(P||Q) &= H(p,q) - H(p) \\ &= \frac{1}{2}\ln|\Sigma_2| - E_p\left\{\ln\left(R_R\frac{\partial R_R}{\partial R_S}\right)\right\} - \ln(\sigma_S\sigma_\theta) \\ &= \frac{1}{2}\ln\left(1+2C_B^2\right) - E_p\left\{\ln\left(\frac{R_R}{R_R|_{\bar{R}_S,\bar{\theta}_R}}\right)\right\} - E_p\left\{\ln\left(\frac{\frac{\partial R_R}{\partial R_S}}{\frac{\partial R_R}{\partial R_S}\Big|_{\bar{R}_S,\bar{\theta}_R}}\right)\right\} \\ &\approx \frac{1}{2}\ln\left(1+2C_B^2\right) \end{aligned} \tag{180}$$

where the final approximation assumes that R_R and $\frac{\partial R_R}{\partial R_S}$ do not vary as a large percent of their value at the mean. The validity of this assumption is examined shortly.

3.2.2 Validity Limits on the KL Divergence Approximation

The validity of the KL divergence approximation is based on the assumption that $E\{\ln R_R\} \approx \ln R_R|_{\bar{R}_S,\bar{\theta}_R}$ and $E\left\{\ln\frac{\partial R_R}{\partial R_S}\right\} \approx \ln\frac{\partial R_R}{\partial R_S}\Big|_{\bar{R}_S,\bar{\theta}_R}$. This is a stronger assumption than $E\{\ln r\} \approx \ln\bar{r}$ in the monostatic case. In particular, for measurements

with extremely high eccentricity the range partial may vary significantly over the support region of the angular error. Let $\epsilon_e = 1 - e$ and $\epsilon_\theta = \bar{\theta}_R + \frac{\pi}{2}$. In this case the approximation is reasonable if

$$\sigma_S \ll \left\{ \bar{R}_S,\ \bar{R}_S \frac{\epsilon_e^2 + \epsilon_\theta^2}{2\epsilon_\theta} \frac{\epsilon_e}{\epsilon_\theta} \right\} \qquad \sigma_\theta \ll \left\{ \frac{\epsilon_e}{\epsilon_\theta},\ \frac{\epsilon_e^2 + \epsilon_\theta^2}{2\epsilon_\theta} \right\} \tag{181}$$

The second part of the restriction on the angle error will most often represent the most stringent restriction, as the restriction on range error is comparable to the monostatic case.

3.3 Gaussian Mixture Modeling for Bistatic Radar

The previous section established an equivalent notion of bias significance which predicts the KL divergence of single-Gaussian fits to a bistatic measurement from the true distribution. In the 2D monostatic case of Chapter 2 the bias significance was also found to be the sole driver of GM KL divergence performance when a mixture of a given size N_G was fitted using ML. This section shows the extension of this property to bistatic radar measurements.

3.3.1 KL Divergence of GM Models for Bistatic Radar

To begin the study of GM models of bistatic measurements, the EM algorithm was used to generate GM fits with a variable number of mixture components to bistatic measurement distributions with C_B varying from 0 to 10. Twenty Monte Carlo samplings of 100k points were made for each bistatic measurement distribution under study. This allows the variability of the outputs from EM to be examined.

Figure 39 shows the KL divergence results from fitting mixtures to a measurement distribution with eccentricity $e = 0.9$ where the transmitter range is 10% of the total range fraction. The angle error is fixed to three different values $\sigma_\theta \in \{10^{-3}, 3 \times 10^{-3}, 10^{-2}\}$ and the bistatic range error σ_S adjusted so that the desired

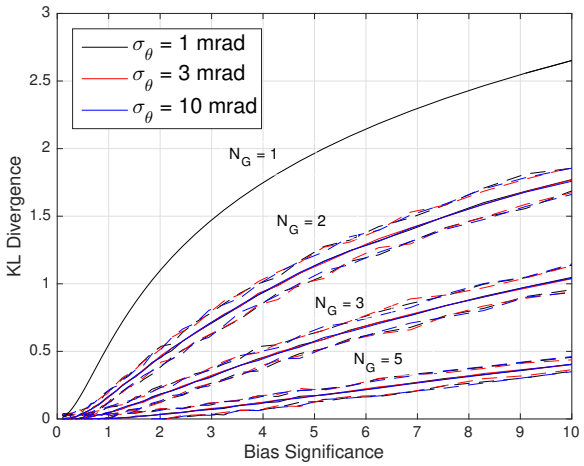

Figure 39: KL Divergence Curves of Mixture Fits to Bistatic Measurement

bias significance is achieved according to (175). The various colors in the plot represent the three different levels of σ_θ. The dashed lines represent the three-sigma bounds on the standard deviation of the sample KL divergence and the solid lines represent the mean of the sample KL divergence. Numbers above the curves indicate the number of mixture components used (1 represents the CB2 model).

These results show that the KL divergence of the GM models are constant (within sampling error bounds) for a given C_B and number of components, despite the changes in σ_θ. Furthermore, by comparison to the results in Figure 3, the values of KL divergence are the same as in the monostatic case. This is significant, as it suggests that the monostatic and bistatic GM fits are closely related, and it should be possible to derive a normalization of the mean and covariance of the ML GM models of bistatic measurements in order to build an efficient lookup table.

These KL divergence results can be reproduced for arbitrary 2D bistatic measurement distributions, provided the restrictions in (181) are met.

3.3.2 Normalized Parameter Definitions

In order to normalize the bistatic GM component means and covariances a similar approach to Chapter 2 is used. First convert the means μ_k to bistatic measurement

space using the inverse measurement transformation (145).

$$\begin{bmatrix} R_{S,k} \\ \theta_{R,k} \end{bmatrix} = \text{C2B}(\mu_k) \tag{182}$$

Next, convert the covariances to the measurement space using linearization around each mean

$$\Sigma_{B,k} = J_{\text{C2B}}(\mu_k)\Sigma_k J_{\text{C2B}}^T(\mu_k) \tag{183}$$

where the Jacobian J_{C2B} is given by (146).

Third, center and scale the bistatic mean parameters according to the errors in the original distribution.

$$\Delta\hat{R}_{S,k} = \frac{R_{S,k} - \bar{R}_S}{\sigma_S C_B} \qquad\qquad \Delta\hat{\theta}_{R,k} = \frac{\theta_{R,k} - \bar{\theta}_R}{\sigma_\theta} \tag{184}$$

Note that these parameters are normalized such that applying the transformation to the single-Gaussian moment matched distribution (CB2) results in $\Delta\hat{R}_{S,k} = -1$ and $\Delta\hat{\theta}_{R,k} = 0$.

Finally, scale the covariance matrix according to the original distribution errors

$$S = \begin{bmatrix} \sigma_S\sqrt{1+2C_B^2} & 0 \\ 0 & \sigma_\theta \end{bmatrix} \qquad\qquad \hat{\Sigma}_{B,k} = S^{-1}\Sigma_{B,k}S^{-T} \tag{185}$$

and decompose this matrix into

$$\hat{\sigma}_{S,k} = \sqrt{[\hat{\Sigma}_{B,k}]_{SS}} \qquad \hat{\sigma}_{\theta,k} = \sqrt{[\hat{\Sigma}_{B,k}]_{\theta\theta}} \qquad \rho_{S\theta,k} = \frac{[\hat{\Sigma}_{B,k}]_{S\theta}}{\hat{\sigma}_{S,k}\hat{\sigma}_{\theta,k}} \tag{186}$$

which represent the normalized errors in each measurement dimension and the correlation coefficient between the errors. In a similar way to the mean parameters, this normalization is designed such that $\hat{\sigma}_{S,k} = 1$, $\hat{\sigma}_{\theta,k} = 1$, and $\rho_{S\theta,k} = 0$ if the transformation is applied to a single-Gaussian moment match.

3.3.3 Measurement Mixture in Terms of Normalized Parameters

Given a bistatic measurement with $R_S \sim N\left(\bar{R}_S, \sigma_S^2\right)$ and $\theta_R \sim N\left(\bar{\theta}_R, \sigma_\theta^2\right)$, a mixture approximation of size N_G is generated using the lookup table of normalized parameters ω_k, $\Delta\hat{R}_{S,k}$, $\Delta\hat{\theta}_{R,k}$, $\hat{\sigma}_{S,k}$, $\hat{\sigma}_{\theta,k}$ and $\rho_{S\theta,k}$ using the following steps.

First, evaluate the bias significance C_B using (175). For the desired number of components N_G, determine values for the normalized parameters evaluated at this C_B based on the lookup table, using interpolation if necessary. Then the component means are given by applying the bistatic measurements transformation (142) in the forward direction.

$$\mu_{B,k} = \begin{bmatrix} \bar{R}_S + \sigma_S C_B \Delta\hat{R}_{S,k} \\ \bar{\theta}_R + \sigma_\theta \Delta\hat{\theta}_{R,k} \end{bmatrix} \qquad \mu_k = \mathrm{B2C}\left(\mu_{B,k}\right) \tag{187}$$

Next, the covariance is found by applying the proper scaling and Jacobian matrix transformation

$$\Sigma_k = T_k \begin{bmatrix} 1 & \rho_{S\theta,k} \\ \rho_{S\theta,k} & 1 \end{bmatrix} T_k^T \qquad T_k = J_{\mathrm{B2C}}(\mu_{B,k}) S \begin{bmatrix} \hat{\sigma}_{S,k} & 0 \\ 0 & \hat{\sigma}_{\theta,k} \end{bmatrix} \tag{188}$$

where S is given in (185). The component weights ω_k are used from the lookup table unmodified.

Note that this is the same procedure that was used in Chapter 2 to specify the monostatic 2D mixtures, except for the fact that the measurement conversion and Jacobian functions are different. Since monostatic radar is simply a special case of bistatic radar, this leads to the significant realization that **the same lookup table can be used to generate both monostatic and bistatic 2D mixtures**. These tables are available in Appendix A. In the appendix, the monostatic column headers $\Delta\hat{r}_k$, $\Delta\hat{\theta}_k$, $\hat{\sigma}_{r,k}$, and $\rho_{r\theta,k}$ give the values of the bistatic parameters $\Delta\hat{R}_{S,k}$, $\Delta\hat{\theta}_{R,k}$, $\hat{\sigma}_{S,k}$, and $\rho_{S\theta,k}$, respectively. See Chapter 2 for a more detailed discussion of the parameter tables.

3.3.4 Mixture Generation Results

Using a lookup table trained on EM algorithm results, mixture representations were generated for a variety of bistatic measurements. The range sum was fixed at $R_S = 100$ km. The baseline L was set by specifying the ellipse eccentricity e, and the receiver mean angle θ_R was set by specifying the receiver range fraction R_{RS}. The angle standard deviation is varied logarithmically over the range $\sigma_\theta = 1$ mrad to $\sigma_\theta = 300$ mrad, and the bias significance varied logarithmically over $C_B = 0.3$ to $C_B = 10$ (which in turn specifies the range standard deviation σ_S). A fixed number of $N_G = 5$ components was used in the fits.

The KL divergence of these fits was then evaluated by sampling 10^6 points from the original measurement distribution and estimating the cross-entropy under the current mixture. These emperical KL divergence quantities are shown in comparison to the predicted KL divergence given in the lookup table.

Figure 40 shows results for $e = 0.4$, $R_{RS} = 0.6$, which represents a scenario where the baseline is a moderate fraction of the bistatic range, and the target is slightly closer to the receiver than the transmitter. The KL divergence matches well for all values for C_B and σ_θ up to approximately $\sigma_\theta = 0.2$, indicating that the generated mixture fits the measurements well. Figure 41 shows a cut of the KL divergence results in Figure 40 at $C_B = 8.34$. Again, the results show that the mixture achieves the expected KL divergence until σ_θ becomes large. The red line plotted in the figure shows the limit indicated by the right hand side of the angle limit in (181). When σ_θ is within an order of magnitude of this limit, performance begins to suffer.

In Figure 42, the same mixture fitting procedure is performed, except the measurement is question has a much higher elliptical eccentricity of $e = 0.9$, indicating that the target is very close to the transmitter-receiver baseline. Additionally, the receiver range fraction is increased to $R_{RS} = 0.9$, which places the target close to the transmitter. The figure shows that KL divergence performance of the mixture model

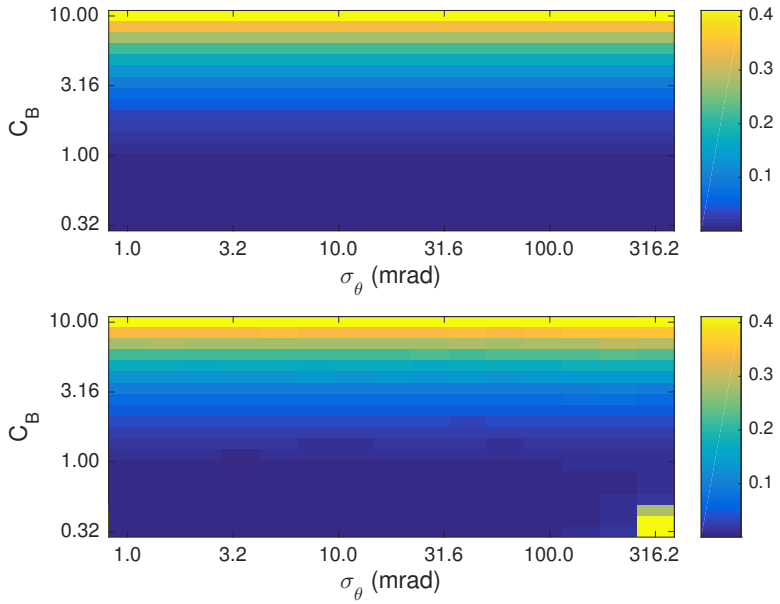

Figure 40: KL Divergence for $e = 0.4$ and $R_{RS} = 0.6$, Predicted (top) and Evaluated (bottom)

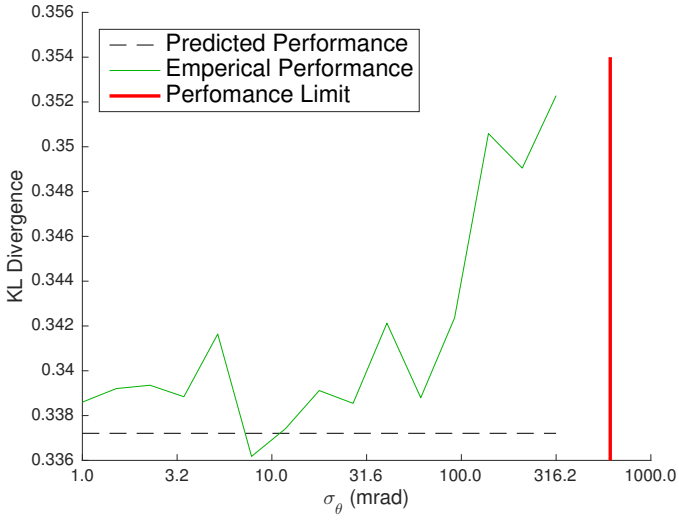

Figure 41: KL Divergence for $C_B = 8.34$, $e = 0.4$ and $R_{RS} = 0.6$

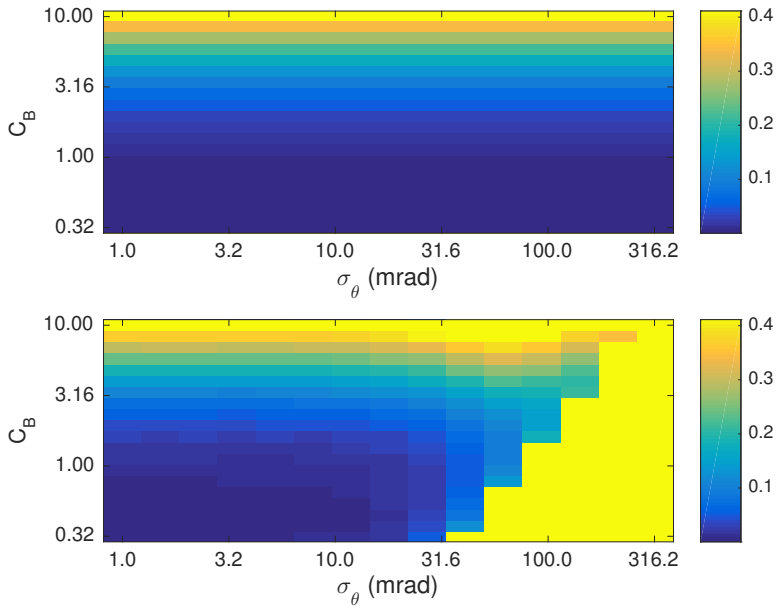

Figure 42: KL Divergence for $e = 0.9$ and $R_{RS} = 0.9$, Predicted (top) and Evaluated (bottom)

is severely degraded for angle error standard deviations above 30 mrad. Figure 43 shows a cut of the KL divergence surface at $C_B = 8.34$. The red line shows that the limit of performance prescribed by (181) is reduced to 100 mrad. As in the previous case, within an order of magnitude of this limit, KL divergence performance of the mixture is poor.

These results demonstrate that the mixture generation procedure for bistatic measurements works well for a variety of bistatic geometries, provided that the limits stated in (181) are respected.

3.4 Filter Implementation

The previous sections demonstrated the effectiveness of the lookup table in modeling bistatic measurements. This table is now implemented into a bistatic version of the MAGMF filter described in Chapter 2. Since no other previous results exist for GM filtering with a bistatic radar, this bistatic MAGMF is compared to number of other single-Gaussian filtering algorithms proposed for bistatic radar tracking. These

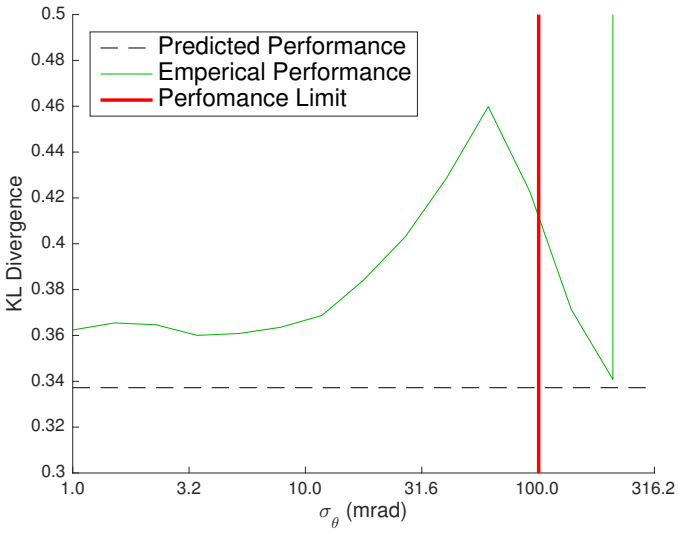

Figure 43: KL Divergence for $C_B = 8.34$, $e = 0.9$ and $R_{RS} = 0.9$

include the standard EKF, a biased converted measurements filter, unbiased converted measurements filter, and UKF. Additionally, an adaptation of the MCAEKF where the bistatic range variance is inflated to limit the bias significance of the bistatic measurements to $C_B \leq 0.2$ is investigated. (It should be noted that an application of MCAEKF to bistatic radar does not currently appear in the literature, so the relative performance of the bistatic adaptation of this filter is informative). Finally, an RPF with 150 k particles and kernel bandwidth $h = 0.5$ s is run as an adaptation of the method used by [24]. These filters are described in detail in Chapter 2. All filters use a nearly-constant-velocity dynamics with the same process noise as in Chapter 2.

The bistatic radar model used in the simulations has baseline $L = 10$ km, angular error $\sigma_\theta = 3$ mrad, and bistatic range error $\sigma_r = 0.1$ m. Measured in the Cartesian coordinate system centered at the midpoint between transmitter and receiver the target begins at $[-4.9\ 1.1]^T$ km, which is very close to the transmitter. The target follows a constant velocity dynamics model with velocity vector of $[-300\ 30]^T$ m/s, so it is moving slowly away from the baseline and towards the transmitter. Based on this geometry and (175), the initial bias significance of the measurements is $C_B = 4.5$ at $t = 0$ s. However, due to the changing position of the target in relation to the

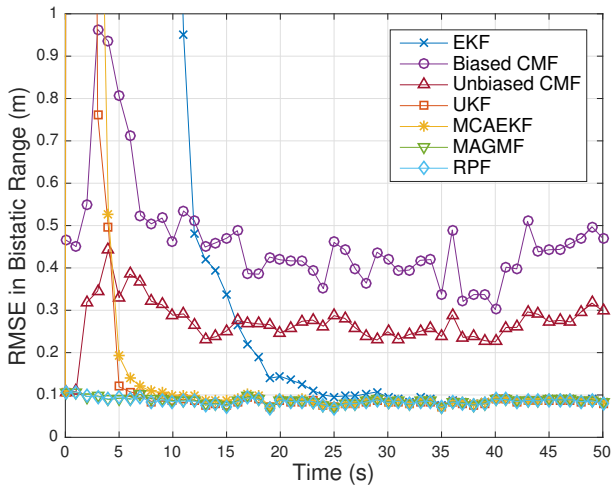

Figure 44: Bistatic Range RMSE for Bistatic Scenario

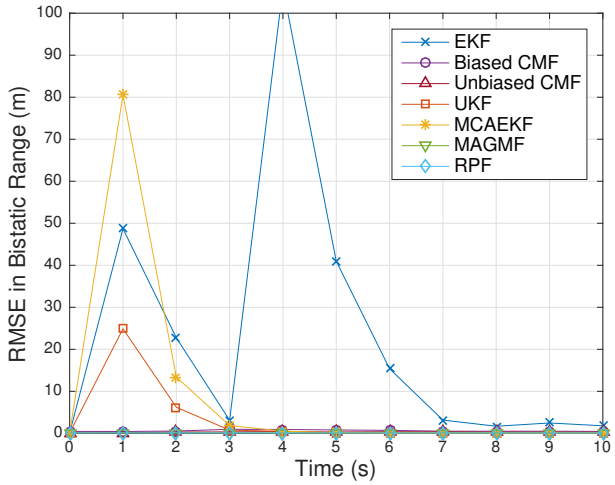

Figure 45: Bistatic Range RMSE for Bistatic Scenario (Wide Zoom)

elements the bias significance decreases to $C_B = 3$ by $t = 50$ s.

Figures 44 through 46 show the position estimation of the filters converted to the original measurement space of bistatic range and receiver angle. As in the monostatic case, the MAGMF and RPF have superior performance in bistatic range early in the track, while the other filters fail to estimate at the same level of precision as the measurement. Notably, the MCAEKF and UKF have extremely poor performance in bistatic range, which is visible when a wider zoom of the performance is shown in Figure 45. Also, the error in receiver angle shown in Figure 46 is much improved for

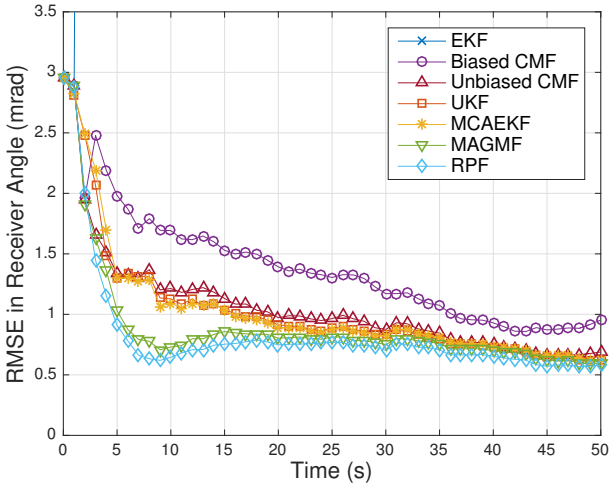

Figure 46: Angle RMSE for Bistatic Scenario

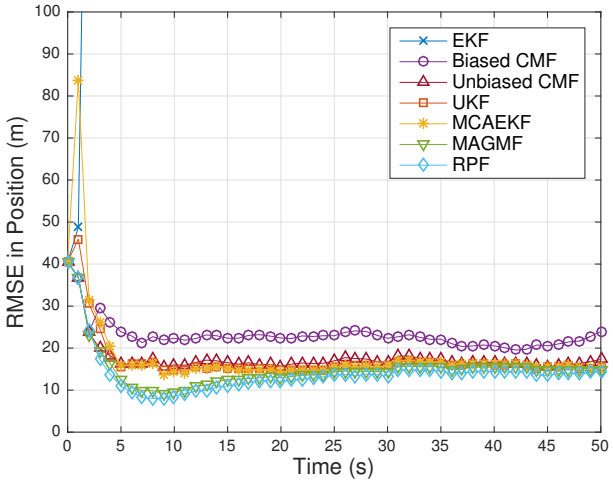

Figure 47: Position RMSE for Bistatic Scenario

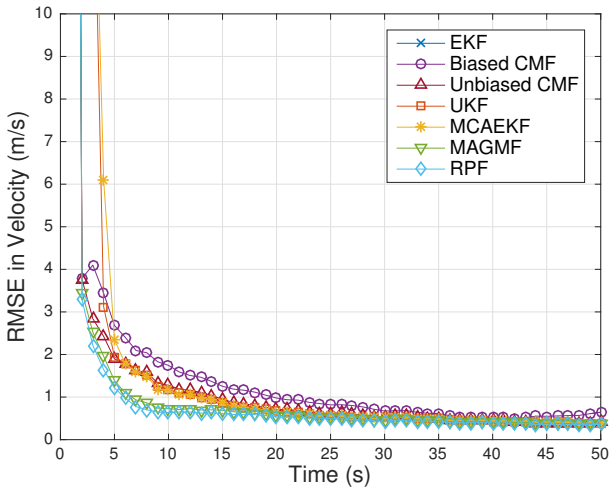

Figure 48: Velocity RMSE for Bistatic Scenario

the RPF and MAGMF compared to the other filters after $t = 5$ sec. (The standard EKF diverges completely in angle so its error is outside the plot axes.) In contrast, in the monostatic case most filters show comparable performance in polar angle RMSE. As a result, the RPF and MAGMF are able to provide significantly better total position RMSE early in the track than the other filters considered. Overall, the RPF slightly outperforms the MAGMF. In velocity RMSE the situation is very similar, as shown in Figure 48. The RPF and MAGMF converge rapidly to an error of less than 1 m/s RMSE after about $t = 5$ sec, which is not achieved by the other filters until after $t = 15$ sec. Again, the RPF slightly outperforms the MAGMF.

Figures 49 and 50 show the NEES consistency metrics for the various filters. Figure 49 evaluates the NEES based on a single-Gaussian moment match of the state position covariance after conversion back to the original measurement space (i.e. the space that would be used for gating), whereas Figure 50 shows the NEES evaluated based on a single-Gaussian moment match to the full state distribution in Cartesian space. The dashed lines represent the 99% confidence region of the NEES based on the number of Monte-Carlo runs performed. All filters are shown to be reasonably consistent except the EKF and the biased CMF, which have NEES that are off the chart. Near the start of the track, the NEES for all filters except the RPF is slightly

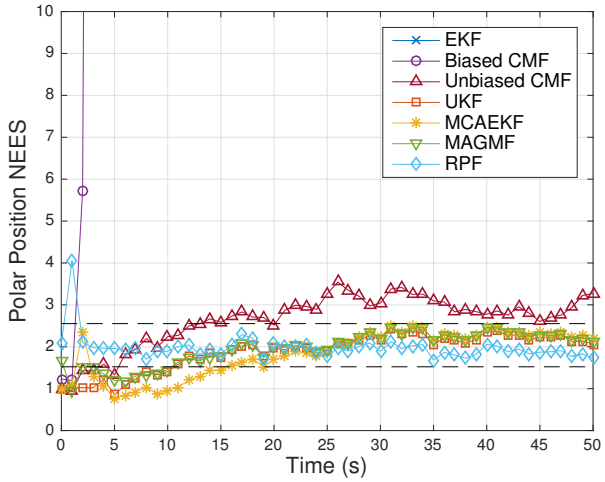

Figure 49: Measurement Space NEES for Bistatic Scenario

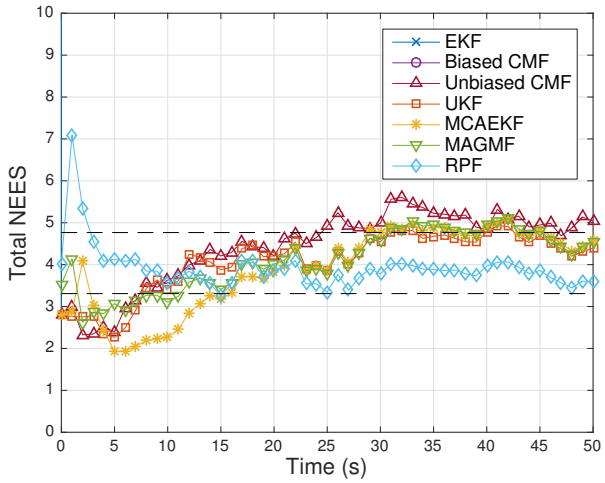

Figure 50: Total Cartesian NEES for Bistatic Scenario

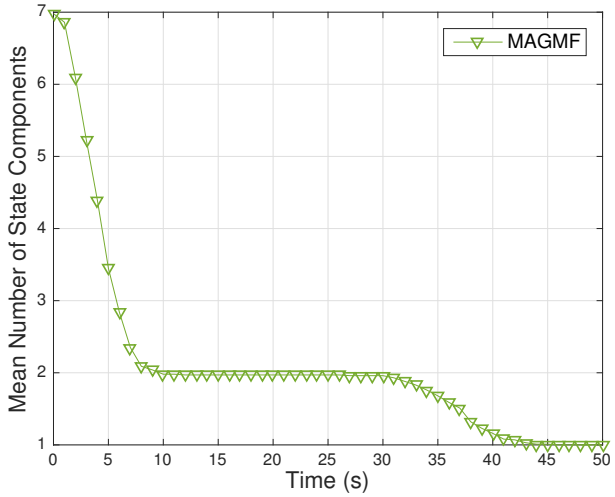

Figure 51: Mean Number of State Components for Bistatic Scenario

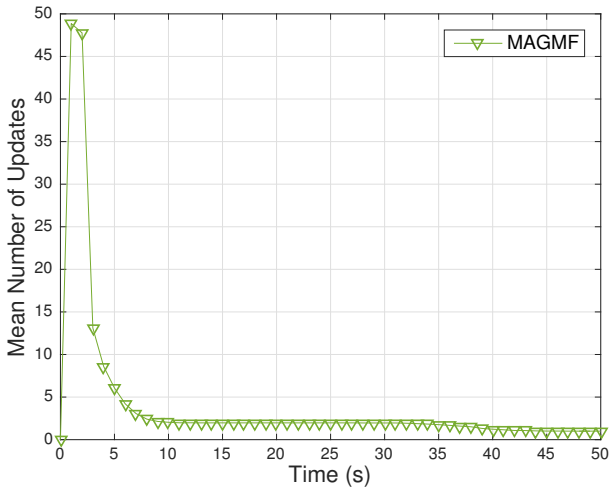

Figure 52: Mean Number of Kalman Updates for Bistatic Scenario

pessimistic due to the use of the one-point initialization strategy which specifies a zero-mean prior distribution with large covariance for the original velocity. The RPF does not use this strategy, but rather uses the differencing method described in [32], therefore is is subject to some artifacts in NEES early in the track.

Finally, the computational complexity of the MAGMF is shown in Figures 51 and 52. Figure 51 shows the number of components used to represent the track state. At the beginning of the track, seven components are used (which is the maximum in the library for this particular implementation), but this quickly decreases as the filter converges with only two or fewer components used after $t = 7$ sec. More significantly, Figure 52 shows the mean number of measurement-to-state Kalman updates needed at each time. Due to the effectiveness of the measurement pre-conditioning strategy in the MAGMF and the rapid decrease in size of the state representation, this cost is only significant for the first couple of updates of the track. By $t = 7$ sec, the MAGMF has a computational cost comparable to that of the single-Gaussian filters. However, the only filter with comparable estimation performance is the RPF, which has a much higher computational cost.

3.5 Conclusions

In this chapter, the flitering methodology developed in Chapter 2 for applying Gaussian mixture modeling to a 2D monostatic radar was successfully extended to a 2D bistatic radar system. First, the bistatic radar transformation was studied and a second-order estimate of the KL divergence of a single-Gaussian fit to a bistatic measurement was derived. Using the relationship of bias significance to KL divergence in the 2D monostatic case, a bistatic bias significance metric was derived which has the same relationship to KL divergence.

The bistatic bias significance was then used to examine the performance of Gaussian mixture fits to bistatic radar measurements. These studies showed that, as in the monostatic case, the KL divergence that is achieved by an ML GM fit to a bistatic measurement is dependent only on its bias significance, rather than on the individual parameters of mean and covariance of the original measurement. This allows the transformations of component means and covariances developed in the 2D monostatic case to be adapted to bistatic measurements. Thus, with a minor change to the normalization procedure, the same lookup tables used to generate 2D monostatic GM models may also be used to generate 2D bistatic GM models.

Using these lookup tables, a bistatic version of the MAGMF from Chapter 2 was implemented and compared to a number of single-Gaussian filtering approaches as well as an RPF. Figure 53 shows the performance of the MAGMF in this scenario in comparison to the other filters tested. The MAGMF has consistency and estimation performance comparable to the RPF while requiring significantly fewer computations. Additionally, the MAGMF adapts the number of GM components as the filter converges, which further reduces computation. None of the single-Gaussian approaches examined were able to match the RMSE performance of the MAGMF. Additionally, the MAGMF provided a more significant benefit to angle RMSE performance than in the monostatic case.

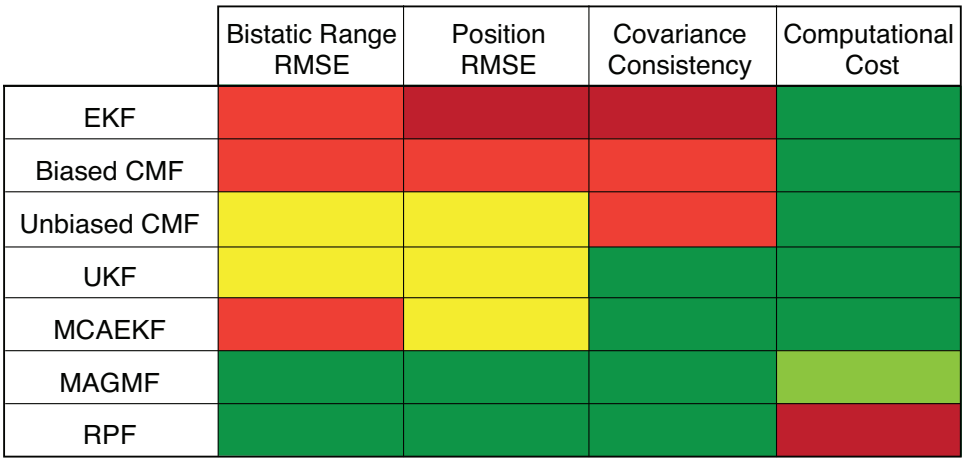

Figure 53: Performance Summary of Filters in 2D Bistatic Simulations

CHAPTER IV

GM FILTERING FOR 3D MONOSTATIC RADARS

Although the study of the 2D contact-lens problem is instructive to provide insight into possible methods that can be used to mitigate the 3D problem, a majority of real-world sensor systems will be concerned with the 3D case. In light of this, many of the previous approaches described in Chapter 2 have been extended to 3D measurements.

The work in [20] and [27] extended the results of [19] to Range-Az-El (RAE) and Range-sine (RUV) coordinates respectively. Additionally, [27] proposed the Modified Covariance Adaptive EKF (MCAEKF), which operates on the principle of inflating the covariance of an RUV measurement so that a consistency criterion is satisfied related to the relative cross-range error of the current track state. This preserves filter consistency, but sacrifices range performance during the period when this inflation occurs.

In order to recover this range performance, 2D Gaussian mixture (GM) solutions were proposed by various authors. Kalman filters using GMs to model the measurement and/or state distribution make use of a GM filter framework developed by Alspach and Sorenson [1]. Tian [28] proposed to split the state PDF into a Gaussian mixture in crossrange in order to satisfy the consistency criteron for the MCAEKF, whereas [21] proposed to split the measurement distribution, but did not discuss an effective method to do so. Previous work by Zhang [32] explored concrete ways to generate these measurement mixture splits, and this method was extended to 3D measurements in [33].

Additionally, [24] explored the use of a Regularized Particle Filter (RPF) to achieve accurate estimation performance in range for 3D measurements. This yielded

promising results, but at the cost of a large number of computations.

In Chapter 2, a comprehensive solution to the contact-lens problem in two dimensions was presented. An information-theoretic approach was used to characterize the Kullback-Leibler (KL) divergence between a single-Gaussian moment matched model of the measurement and the true distribution, as well as between Gaussian mixture (GM) models generated using the Expectation Maximization (EM) algorithm. It was found that the bias significance parameter C_B can be directly linked to the KL divergence performance of both single-Gaussian and GM models of the measurement, and this fact was used to build a lookup table of GM fits to polar radar measurements of arbitrary mean and variance based on the results of EM. Combined with a novel measurement conditioning procedure that temporarily relocates the support of the measurement to be local to the support of the prior state, this lookup table approach was used to implement the Measurement-Adaptive Gaussian Mixture Filter (MAGMF) in two dimensions. This filter was able to achieve excellent range estimation performance and covariance consistency with a more limited number of computations than previous GM approaches and very much fewer than a particle filter.

The goal of this chapter is to extend the theory and techniques used to implement the MAGMF in 2D to the 3D RAE case. This chapter is organized as follows. Firstly, the RAE transformation is presented and studied in order to derive KL divergence expressions for the single-Gaussian moment matched models compared to the true distribution. This is used to extend the definition of bias significance to 3D measurements. Based on these expressions, the EM algorithm is used to generate GM fits of varying size to RAE measurements with common bias significance. It is found that, unlike the 2D case, in the 3D the ratio of Az and El errors within a given measurement materially affects the KL divergence that may be achieved by a GM fit as opposed to simply the bias significance.

Due to this additional degree of freedom, further study in this chapter is focused primarily on the case where Az and El error have equal effect on the bias significance. The GM solutions generated by the EM algorithm for these cases are studied and categorized by geometric regularity and KL divergence performance. A transformation procedure similar to that of Chapter 2 is then developed to allow the parameters of the means and covariances of these mixtures to be stored in a lookup table of limited size. Combined with the conditioning method developed in Chapter 2, this lookup table is then used to create a 3D implementation of the MAGMF algorithm. This algorithm is then run against several long-range radar scenarios in comparison to several of the existing techniques described above and results of this new filter presented.

4.1 Range-Az-El Measurements Transformation

The coordinate conversion for three-dimensional measurements given in terms of bearing angles is given by

$$x = r\cos\alpha\cos\varepsilon \qquad y = r\sin\alpha\cos\varepsilon \qquad z = r\sin\varepsilon \tag{189}$$

For convenience, this can be defined as a vector function. With $x_S = \begin{bmatrix} r & \alpha & \varepsilon \end{bmatrix}^T$ and $x_C = \begin{bmatrix} x & y & z \end{bmatrix}^T$

$$x_C = \mathrm{S2C}\left(x_S\right) = \begin{bmatrix} r\cos\alpha\cos\varepsilon \\ r\sin\alpha\cos\varepsilon \\ r\sin\varepsilon \end{bmatrix} \tag{190}$$

The transform Jacobian is

$$\frac{\partial}{\partial x_S}\mathrm{S2C} = J_{\mathrm{S2C}}(x_S) = \begin{bmatrix} \cos\alpha\cos\varepsilon & -r\sin\alpha\cos\varepsilon & -r\cos\alpha\sin\varepsilon \\ \sin\alpha\cos\varepsilon & r\cos\alpha\cos\varepsilon & -r\sin\alpha\sin\varepsilon \\ \sin\varepsilon & 0 & r\cos\varepsilon \end{bmatrix} \tag{191}$$

with determinant

$$|J_{\mathrm{S2C}}| = r^2 \cos\varepsilon \tag{192}$$

The inverse transformation and Jacobian are

$$x_S = \mathrm{C2S}\left(x_C\right) = \begin{bmatrix} \sqrt{x^2+y^2+z^2} \\ \tan^{-1}\frac{x}{y} \\ \sin^{-1}\frac{z}{r} \end{bmatrix} \tag{193}$$

$$J_{\mathrm{C2S}}(x_C) = J^{-1}_{\mathrm{S2C}}(x_C) = \begin{bmatrix} \frac{x}{r} & \frac{y}{r} & \frac{z}{r} \\ -\frac{y}{r_{xy}} & \frac{x}{r_{xy}} & 0 \\ -\frac{xz}{r^2 r_{xy}} & -\frac{yz}{r^2 r_{xy}} & \frac{r_{xy}}{r^2} \end{bmatrix} \tag{194}$$

where $r_{xy} = \sqrt{x^2+y^2}$.

Given the measurement statistics

$$r \sim N\left(\bar{r}, \sigma_r^2\right) \qquad \alpha \sim N\left(\bar{\alpha}, \sigma_\alpha^2\right) \qquad \varepsilon \sim N\left(\bar{\varepsilon}, \sigma_\varepsilon^2\right) \tag{195}$$

let $\Delta r = r - \bar{r}$, $\Delta\alpha = \alpha - \bar{\alpha}$, $\Delta\varepsilon = \varepsilon - \bar{\varepsilon}$ be the zero mean errors of the measurements. Then $\bar{x}_S = E\left\{x_S\right\}$ and $\Delta x_S = x_S - \bar{x}_S$ the vectorized mean and zero mean errors. Define the rotations

$$R_z(\bar{x}_S) = \begin{bmatrix} \cos\bar{\alpha} & \sin\bar{\alpha} & 0 \\ -\sin\bar{\alpha} & \cos\bar{\alpha} & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{196}$$

$$R_y(\bar{x}_S) = \begin{bmatrix} \cos\bar{\varepsilon} & 0 & \sin\bar{\varepsilon} \\ 0 & 1 & 0 \\ -\sin\bar{\varepsilon} & 0 & \cos\bar{\varepsilon} \end{bmatrix} \tag{197}$$

Then the Cartesian coordinates of the converted measurements can be expressed in the rotated system

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = R_y R_z \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{198}$$

Note that

$$R_z^T R_y^T = J_{\mathrm{S2C}}(\bar{x}_S) \begin{bmatrix} 1 & 0 & 0 \\ 0 & \bar{r}\cos\bar{\varepsilon} & 0 \\ 0 & 0 & \bar{r} \end{bmatrix}^{-1} \tag{199}$$

so that $R_y R_z$ is equivalent to a unitary transformation into the normalized version of the axes described by the columns of the Jacobian matrix evaluated at the mean. This transform can be expanded to

$$x' = (\bar{r} + \Delta r)\left[\cos\Delta\varepsilon + (\cos\Delta\alpha - 1)\cos\bar{\varepsilon}\cos\varepsilon\right] \tag{200}$$

$$y' = (\bar{r} + \Delta r)\sin\Delta\alpha\cos\varepsilon \tag{201}$$

$$z' = (\bar{r} + \Delta r)\left[\sin\Delta\varepsilon - (\cos\Delta\alpha - 1)\sin\bar{\varepsilon}\cos\varepsilon\right] \tag{202}$$

Taking the expectation

$$\mu' = E\left\{\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}\right\} = \begin{bmatrix} \bar{r}E\left\{\cos\Delta\varepsilon\right\}\left(1 + \cos^2\bar{\varepsilon}\left(E\left\{\cos\Delta\alpha\right\} - 1\right)\right) \\ 0 \\ -\bar{r}E\left\{\cos\Delta\varepsilon\right\}\sin\bar{\varepsilon}\cos\bar{\varepsilon}\left(E\left\{\cos\Delta\alpha\right\} - 1\right) \end{bmatrix} \tag{203}$$

then the mean in the original coordinates is given by $\mu = R_z^T R_y^T \mu'$. Now, $E\left\{\cos\Delta\alpha\right\}$ and $E\left\{\cos\Delta\varepsilon\right\}$ can be approximated to first order as 1 and to second order as $1 - \frac{\Delta\alpha^2}{2}$ and $1 - \frac{\Delta\varepsilon^2}{2}$, respectively. Substituting these terms into the expectation gives $\mu_1 = R_z^T R_y^T \mu_1'$ as a first order approximation of the mean, where $\mu_1' = \begin{bmatrix} \bar{r} & 0 & 0 \end{bmatrix}^T$ and $\mu_2 = R_z^T R_y^T \mu_2'$ as a second order approximation of the mean, with

$$\mu_2' = \begin{bmatrix} \bar{r}\left(1 - \frac{\sigma_\varepsilon^2}{2} - \frac{\sigma_\alpha^2}{2}\cos^2\bar{\varepsilon}\right) \\ 0 \\ \bar{r}\frac{\sigma_\alpha^2}{2}\sin\bar{\varepsilon}\cos\bar{\varepsilon} \end{bmatrix} \tag{204}$$

Note that terms involving the product of $\Delta\alpha^2$ and $\Delta\varepsilon^2$ have been removed in the second order estimate.

For the covariance, first subtract the mean from each coordinate

$$\begin{aligned} x' - E\{x'\} = \bar{r}\Big[& -\cos\bar{\varepsilon}\sin\bar{\varepsilon}\,(\cos\Delta\alpha - 1)\sin\Delta\varepsilon \\ & + \cos^2\bar{\varepsilon}\big(\cos\Delta\alpha\cos\Delta\varepsilon \\ & \qquad - E\{\cos\Delta\alpha\}\,E\{\cos\Delta\varepsilon\}\big) \\ & + \big(1 - \cos^2\bar{\varepsilon}\big) \\ & \qquad (\cos\Delta\varepsilon - E\{\cos\Delta\varepsilon\})\Big] \\ & + \Delta r\,[\cos\Delta\varepsilon + (\cos\Delta\alpha - 1)\cos\bar{\varepsilon}\cos\varepsilon] \end{aligned} \tag{205}$$

$$y' - E\{y'\} = (\bar{r} + \Delta r)\sin\Delta\alpha\cos\varepsilon \tag{206}$$

$$\begin{aligned} z' - E\{z'\} = & \bar{r}\cos\bar{\varepsilon}\sin\bar{\varepsilon}\Big[(\cos\Delta\varepsilon - E\{\cos\Delta\varepsilon\}) \\ & - \big(\cos\Delta\alpha\cos\Delta\varepsilon \\ & \qquad - E\{\cos\Delta\alpha\}\,E\{\cos\Delta\varepsilon\}\big)\Big] \\ & + \bar{r}\sin\Delta\varepsilon \\ & + \Delta r\,[\sin\Delta\varepsilon - (\cos\Delta\alpha - 1)\sin\bar{\varepsilon}\cos\varepsilon] \end{aligned} \tag{207}$$

Now, let $\Delta x_1'$, $\Delta y_1'$, and $\Delta z_1'$ be first order approximations of $x' - E\{x'\}$, $y' - E\{z'\}$, and $z' - E\{z'\}$, respectively. Similarly, let let $\Delta x_2'$, $\Delta y_2'$, and $\Delta z_2'$ be second order approximations of these quantities. These approximations are performed by substituting $\sin\Delta\alpha$, $\cos\Delta\alpha$, $\sin\Delta\varepsilon$, and $\cos\Delta\varepsilon$ with their respective first or second order Taylor series expansions.

$$\Delta x_1' = \Delta r$$

$$\Delta x_2' = \Delta r + \bar{r}\Big[\cos^2\bar{\varepsilon}\left(\frac{\sigma_\alpha^2}{2} - \frac{\Delta\alpha^2}{2}\right) + \left(\frac{\sigma_\varepsilon^2}{2} - \frac{\Delta\varepsilon^2}{2}\right)\Big]$$

$$\Delta y_1' = \bar{r}\Delta\alpha\cos\bar{\varepsilon}$$

$$\Delta y_2' = (\bar{r} + \Delta r)\,\Delta\alpha\cos\bar{\varepsilon}$$

$$\Delta z_1' = \bar{r}\Delta\varepsilon$$

$$\Delta z_2' = (\bar{r} + \Delta r)\,\Delta\varepsilon - \bar{r}\cos\bar{\varepsilon}\sin\bar{\varepsilon}\left(\frac{\sigma_\alpha^2}{2} - \frac{\Delta\alpha^2}{2}\right)$$

Using these expressions, form first and second order approximations of the transformed covariance. Let

$$\Sigma_1' = E\left\{ \begin{bmatrix} \Delta x_1' \\ \Delta y_1' \\ \Delta z_1' \end{bmatrix} \begin{bmatrix} \Delta x_1' \\ \Delta y_1' \\ \Delta z_1' \end{bmatrix}^T \right\} = \begin{bmatrix} \sigma_r^2 & 0 & 0 \\ 0 & \bar{r}^2\sigma_\alpha^2 \cos^2\bar{\varepsilon} & 0 \\ 0 & 0 & \bar{r}^2\sigma_\varepsilon^2 \end{bmatrix} \tag{208}$$

and

$$\Sigma_2' = E\left\{ \begin{bmatrix} \Delta x_2' \\ \Delta y_2' \\ \Delta z_2' \end{bmatrix} \begin{bmatrix} \Delta x_2' \\ \Delta y_2' \\ \Delta z_2' \end{bmatrix}^T \right\} = \begin{bmatrix} \Sigma_{xx,2}' & 0 & \Sigma_{xz,2}' \\ 0 & \Sigma_{yy,2}' & 0 \\ \Sigma_{xz,2}' & 0 & \Sigma_{zz,2}' \end{bmatrix} \tag{209}$$

where

$$\Sigma_{xx,2}' = \sigma_r^2 + \bar{r}^2\left(\frac{\sigma_\alpha^4}{2}\cos^4\bar{\varepsilon} + \frac{\sigma_\varepsilon^4}{2}\right)$$

$$\Sigma_{yy,2}' = \left(\bar{r}^2 + \sigma_r^2\right)\sigma_\alpha^2\cos^2\bar{\varepsilon}$$

$$\Sigma_{xz,2}' = -\bar{r}^2\frac{\sigma_\alpha^4}{2}\cos^3\bar{\varepsilon}\sin\bar{\varepsilon}$$

$$\Sigma_{zz,2}' = \left(\bar{r}^2 + \sigma_r^2\right)\sigma_\varepsilon^2 + \bar{r}^2\frac{\sigma_\alpha^4}{2}\cos^2\bar{\varepsilon}\sin^2\bar{\varepsilon}$$

In order to find approximations of the covariance of the original vector $\begin{bmatrix} x & y & z \end{bmatrix}^T$, Σ_1' and Σ_2' may simply be transformed by the inverse of the original rotation

$$\Sigma_1 = R_z^T R_y^T \Sigma_1' R_y R_z \tag{210}$$

$$\Sigma_2 = R_z^T R_y^T \Sigma_2' R_y R_z \tag{211}$$

In [10] it was shown that the KL divergence of a moment matched Gaussian approximation of the converted measurement distribution from the true distribution is dependent primarily on the determinant of the covariance matrix. For this purpose, the second order estimate of the covariance matrix will suffice. As $R_y R_z$ is a unitary transformation, it is true that $|\Sigma_2| = |\Sigma_2'|$, where the latter is much easier to compute.

$$|\Sigma_2'| = \Sigma_{yy,2}'\left(\Sigma_{xx,2}'\Sigma_{zz,2}' - \left(\Sigma_{xz,2}'\right)^2\right) \tag{212}$$

Let

$$C_{B,\varepsilon} = \bar{r}\frac{\sigma_{\bar{\varepsilon}}^2}{2\sigma_r} \tag{213}$$

be defined as the "elevation bias significance" and

$$C_{B,\alpha} = \bar{r}\frac{\sigma_{\alpha}^2 \cos^2\bar{\varepsilon}}{2\sigma_r} \tag{214}$$

be defined as the "azimuth bias significance" in analogy to the 2D case presented in [9]. Then

$$\begin{aligned}\Sigma'_{xx,2}\Sigma'_{zz,2} - \left(\Sigma'_{xz,2}\right)^2 = \left(\bar{r}^2 + \sigma_r^2\right)\sigma_{\varepsilon}^2\sigma_r^2\left(1 + 2C_{B,\varepsilon}^2 + 2C_{B,\alpha}^2\right) \\ + 2\sigma_r^4 C_{B,\alpha}^2 \tan^2\bar{\varepsilon}\left(1 + C_{B,\varepsilon}^2\right)\end{aligned} \tag{215}$$

Assuming $\sigma_r \ll \bar{r}$ and $\sigma_{\varepsilon} << 1$, this may be approximated as

$$\Sigma'_{xx,2}\Sigma'_{zz,2} - \left(\Sigma'_{xz,2}\right)^2 \approx \bar{r}^2\sigma_{\varepsilon}^2\sigma_r^2\left(1 + 2C_{B,\varepsilon}^2 + 2C_{B,\alpha}^2\right) \tag{216}$$

so that the estimate of covariance determinant is

$$|\Sigma'_2| \approx \bar{r}^4\cos^2\bar{\varepsilon}\left(1 + 2C_{B,\varepsilon}^2 + 2C_{B,\alpha}^2\right)\sigma_r^2\sigma_{\varepsilon}^2\sigma_{\alpha}^2 \tag{217}$$

The first order estimate is given by

$$|\Sigma'_1| = \bar{r}^4\cos^2\bar{\varepsilon}\,\sigma_r^2\sigma_{\varepsilon}^2\sigma_{\alpha}^2 \tag{218}$$

In Chapter 2 it was shown that if x is Gaussian with distribution $p(x)$ and $y = f(x)$ is modeled by distribution $q(y)$ the the cross-entropy from q to p is given by

$$H(p,q) = -E_p\left\{\log q_y\left(f\left(x\right)\right)\right\} - E_p\left\{\log\left|\frac{\partial f}{\partial x}\right|\right\} \tag{219}$$

Using this result, the cross-entropy of the moment matched Gaussian distribution is given by

$$\begin{aligned}H(p,q) &= \frac{3}{2}\left(1 + \ln(2\pi)\right) + \frac{1}{2}\ln|\Sigma_q| - E\left\{\ln\left(r^2\cos\varepsilon\right)\right\} \\ &\approx \frac{3}{2}\left(1 + \ln(2\pi)\right) + \frac{1}{2}\ln|\Sigma'_2| - \ln\left(\bar{r}^2\cos\bar{\varepsilon}\right)\end{aligned} \tag{220}$$

The entropy of the original distribution is

$$H(p) = \frac{3}{2}\left(1 + \ln(2\pi)\right) + \ln\left(\sigma_r \sigma_\alpha \sigma_\varepsilon\right) \tag{221}$$

Therefore, the estimate of KL divergence of the moment matched Gaussian distribution from the true measurement is

$$\begin{aligned} D_{KL}\left(P||Q\right) &= H\left(p,q\right) - H\left(p\right) \\ &\approx \frac{1}{2}\ln\left(1 + 2C_{B,\varepsilon}^2 + 2C_{B,\alpha}^2\right) \end{aligned} \tag{222}$$

Note that using the definitions of azimuth and elevation bias significance, (204) may be re-written as

$$\mu_2' = \begin{bmatrix} \bar{r} - \sigma_r(C_{B,\alpha} + C_{B,\varepsilon}) \\ 0 \\ \sigma_r C_{B,\alpha} \tan\bar{\varepsilon} \end{bmatrix} \tag{223}$$

where the top (x') term dominates the total range, so it is seen that these quantities indeed describe the magnitude of the range bias in the mean, as well as the KL divergence and covariance inflation. To simplify the expression of the KL divergence, define the "total bias significance" as

$$C_B = \sqrt{C_{B,\varepsilon}^2 + C_{B,\alpha}^2} \tag{224}$$

so that the KL divergence estimate is given by

$$D_{KL}\left(P||Q\right) \approx \frac{1}{2}\ln\left(1 + 2C_B^2\right) \tag{225}$$

in analogy to the 2D case described in [9]. Furthermore, define the quantity

$$R_B = \frac{C_{B,\alpha}}{C_{B,\varepsilon}} = \frac{\sigma_\alpha^2 \cos^2\bar{\varepsilon}}{\sigma_{\bar{\varepsilon}}^2} \tag{226}$$

as the "bias significance ratio". It should be noted that the quantity $\sigma_r C_B$ does not describe the total bias in range, so that the name "bias significance" is somewhat

misleading in this case. However, incorporating the bias significance ratio and using the definition of total bias significance results in

$$[\mu_2']_r - \bar{r} = \Delta r = -\sigma_r C_B \frac{1 + R_B}{\sqrt{1 + R_B^2}} \tag{227}$$

which for the $R_B = 1$ case reduces to $\Delta r = -\sqrt{2}\sigma_r C_B$.

4.2 ML Fitting of Gaussian Mixtures to 3D Measurements

As in Chapter 2, the EM algorithm was used to perform ML fits of GM parameters to the 3D converted measurements. As an initial experiment, Monte Carlo runs of the EM algorithm were performed to fit five Gaussian components ($N_G = 5$) to sets of measurements with a fixed total bias significance of $C_B = 5$. The measurement range was fixed to $\bar{r} = 100$ km and the range standard deviation to $\sigma_r = 0.01$ m. The elevation angle ε was varied from $+90$ to -90 degrees and the ratio of angle errors $\sqrt{R_B} = \frac{\sigma_\alpha \cos\varepsilon}{\sigma_\varepsilon}$ of the measurements was allowed to vary logarithmically from 0.1 to 10 while keeping the total bias significance constant.

Five Monte Carlo runs of EM were performed for each angle error ratio and elevation point in the grid on using $100,000$ samples of the converted measurement distribution. Over these five runs, the minimum KL divergence solution was recorded. These KL divergence results are shown in Figure 54. The log of the Az-El error ratio is plotted in order to provide equal treatment to reciprocal cases of the ratio (e.g. 0.1 vs 10). The first main observation from these results is that the KL divergence achieved by the mixtures is essentially independent of the elevation angle (with some exception near the singularity at $\varepsilon = \pi/2$). In light of this, the results for each individual elevation sample are replotted on top of each other in Figure 55, which allows the KL divergence values to more easily be seen.

Figure 55 highlights the second major observation that the KL divergence that can be achieved for an unity error ratio ($R_B = 1$) is much higher than in the case when either Az or El error dominates independently. Indeed, the performance for a

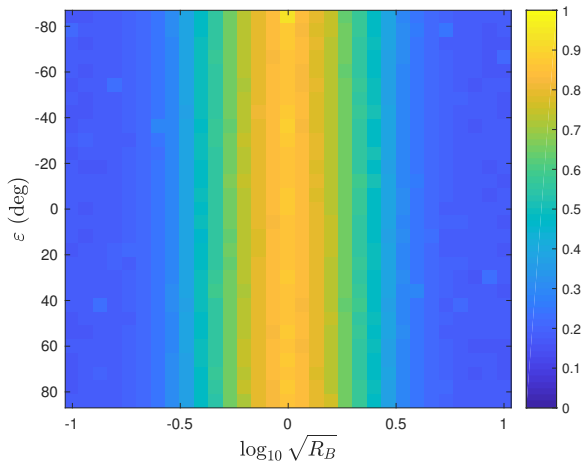

Figure 54: 3D KL Divergence of Mixture Fits with $C_B = 5$ and $N_G = 5$

large or small error ratio is similar to the 2D case. Figure 56 shows the KL divergence results from Chapter 2 for 2D GM fits, with a circular marker added for the point under consideration of $C_B = 5$, $N_G = 5$. The 2D KL divergence is around 0.173, which is comparable to the 3D performance at the ends of the plot in Figure 55 (error ratio 0.1 or 10). This makes sense, as measurements with a large or small error ratio can essentially be modeled as a trivial extension of the 2D case. However, in the unity error ratio case, the KL divergence is 0.85, over four times the 2D value. Therefore, more components are needed in the $R_B = 1$ case to ensure good modeling performance.

These results stand in contrast to the 2D case. In both 2D and 3D, measurements with the same bias significance have the KL divergence under a single Gaussian fit. Additionally, in 2D, these measurements have the same KL divergence under a Gaussian mixture fit of given size (N_G). However, in 3D, the ratio of Az to El error within the bias significance will change the KL divergence of a GM fit to the measurement. Therefore, a lookup-table procedure for modeling 3D measurements with a Gaussian mixture must be sensitive to the error ratio.

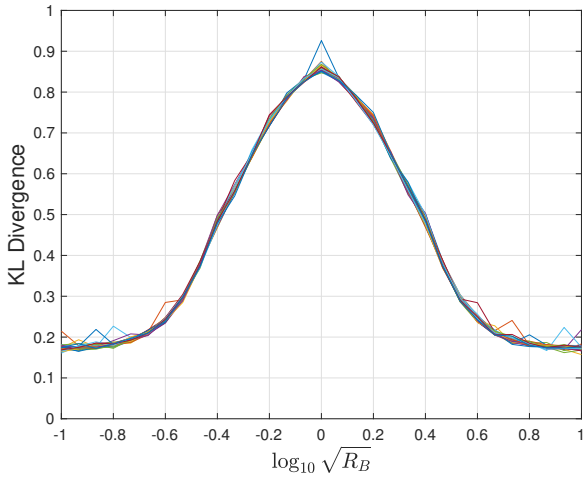

Figure 55: 3D KL Divergence of Mixture Fits with $C_B = 5$ and $N_G = 5$ - Elevation Cuts

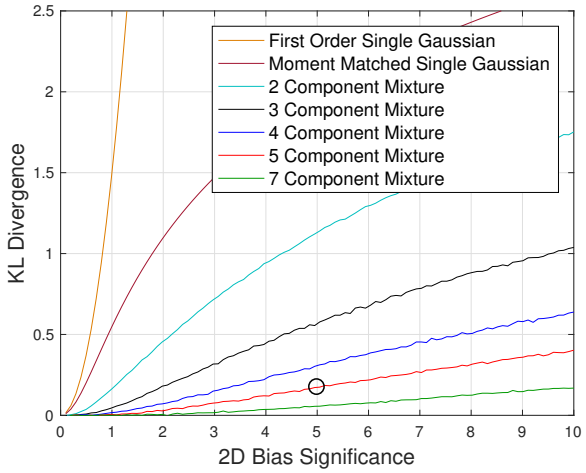

Figure 56: 2D KL Divergence of Mixture Fits

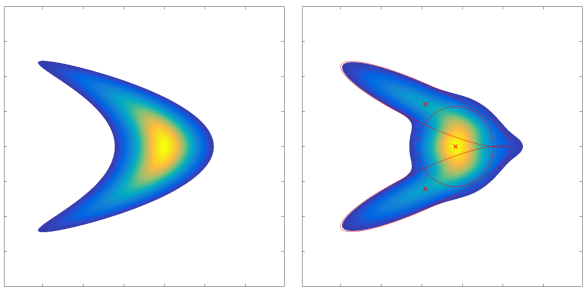

Figure 57: 2D Measurement Distribution with $C_B = 1$ (left) and $N_G = 3$ Component Mixture Fit (right)

4.3 Geometries of 3D Gaussian Mixture Fits

As the EM algorithm finds locally optimum solutions of the mixture parameters with respect to the KL divergence, it is useful to initialize the algorithm with an arrangement of components reasonably close to the desired solution. This both improves the rate of convergence and ensures that the globally optimal solution is reached (or potentially, a particular locally optimal solution with additional desirable properties such as geometric regularity). Figure 57 shows the layout of mixture components for a GM fit with $N_G = 3$ to a 2D converted measurement distribution (right side). The components are arranged with their means positioned approximately on the circular iso-range contour of the mean range at semi-regular intervals. This arrangement extends to higher numbers of components, so that the cross-range support of the distribution is simply partitioned into greater and greater numbers of pieces by the GM. Therefore, a simple, effective strategy for initializing EM to find these 2D solutions is to place the component means with equal spacing along the crossrange support as described in [28].

However, in 3D the geometric arrangement of component means is not as straightforward, especially in the case of unity Az-El bias significance ratio ($R_B = 1$). For many choices of N_G, multiple sub-optimal solutions exist. Furthermore, a certain geometry may be optimal for low bias significance, but sub-optimal for a higher bias significance.

Due to the reduced KL divergence performance of mixtures fitted to a measurement with unity Az-El bias significance ratio compared to the 2D case, it is necessary to use a larger number of components to represent the measurement to ensure good performance. Storing the parameters of a large number of components in a lookup table may prove inconvenient if they are not arranged in a regular geometry. Therefore, it is desired to choose the numbers of components N_G that admit a solution having a regular geometry. Furthermore, a slightly sub-optimal local optimum solution may be preferred to the globally optimal arrangement for a given N_G if the solution is regular.

In order to study the geometric arrangement of 3D GM solution components, we first consider the $R_B = 1$ case. The EM algorithm with random initialization was used to fit a GM of given number of components N_G to a converted measurement with $\bar{r} = 100$ km, $\bar{\varepsilon} = \bar{\alpha} = 0$, and $\sigma_\alpha = \sigma_\varepsilon = 1$ mrad. The range standard deviation σ_r was chosen using (224) such that the measurement has a given total bias significance C_B. The range of bias significances studied for each N_G was restricted to those resulting in a mixture fit with KL divergence approximately in the interval $(0.05, 0.2)$, as it was found in the 2D case [9] that values outside this interval are less useful for application to a Kalman filter.

The results of this case are plotted in the following manner. Firstly, the transformation (198) is applied in order to express the coordinates in terms of the axes in the columns of the transform Jacobian matrix. Rather than use x',y', and z' to express the coordinates in this context, relabel the axes $\partial R = x'$, $\partial Az = y'$, and $\partial El = z'$ in order to more straightforwardly identify them with the columns of the Jacobian to which they correspond. Furthermore, the $\Delta\partial R$ axes refers to ∂R with the mean range $\bar{r}$ subtracted. Finally, in order to graphically represent the PDF results, the full distributions are marginalized into the $\partial Az, \partial El$ plane and the $\Delta\partial R, \partial Az$ plane, and these marginal PDFs are plotted in a scaled colormap over a grid of appropriate size

around the mean. In the case that a mixture PDF is represented, the means of the components are additionally represented by a red cross, and the 2-sigma covariances (contours of Mahalanobis distance 4) by a red ellipse.

4.3.1 Geometry Description Codes

Many of the GM solution geometries exhibit similar characteristics when viewed in the Az/El plane. These features include one or more rings of components with similar weight arranged with their means in a circle, a central component covering the mean range, az, and el, and/or a single component with large covariance and low weight which has a large offset in range in order to provide coverage of the tails of the distribution. Therefore, it is useful to develop a common notation for these geometries.

If a given geometry contains a central component (centered in range, az, and el), this will be denoted by the letter 'C'. The presence of the low weight range tail component is denoted by 'T'. A set of similar components resembling a ring in the Az/El plane is denoted by the number of components in the set, followed by an 'R' for ring. To describe the whole geometry, the component features will be listed in order of increasing radius of their mean position in the Az/El plane, and individual feature codes separated by hyphens. The only exception is that the range tail 'T' is always listed last.

For example, the $N_G = 5$ solution shown in Figure 68 with a central component, a ring of 3 components surrounding, and a range tail is denoted by "C-3R-T", whereas the $N_G = 9$ solution with an inner ring of three components surrounded by an outer ring of six is "3R-6R."

4.3.2 $N_G = 2$ Solutions

For the $N_G = 2$ case, a measurement with bias significance of $C_B = 0.7$ was fitted. The true marginals of this measurement are shown in Figure 58. Two main regular

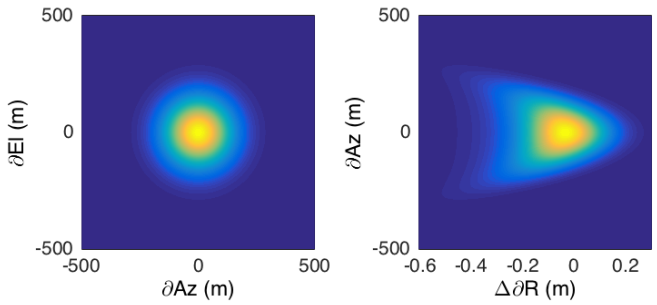

Figure 58: True Marginal PDFs for Az-El (left) and Range-Az (right) for $C_B = 0.7$

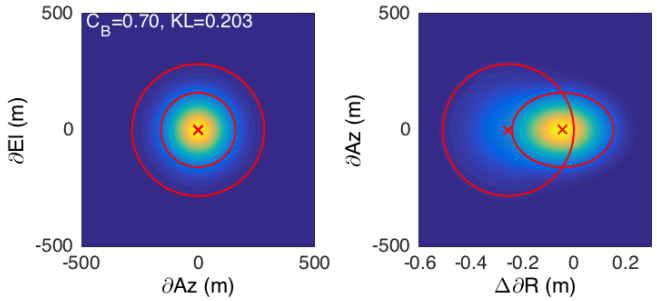

Figure 59: GM Marginal PDFs with $N_G = 2$ C-T for Az-El (left) and Range-Az (right) for $C_B = 0.7$

geometries result for the mixtures, one with a range tail (C-T) shown in Figure 59 and one without (2R) in Figure 60. Both geometries achieve a KL divergence of approximately 0.2, which is adequate for modeling in a Kalman filter. The C-T geometry has approximately 10% better KL divergence performance for the moderate bias significance studied. However, 2R does a slightly better job of representing the structure of the tails of the distribution.

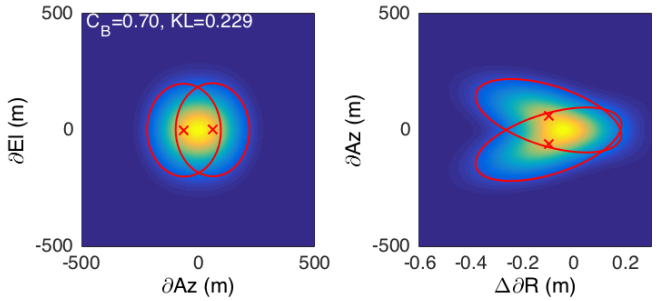

Figure 60: GM Marginal PDFs with $N_G = 2$ 2R for Az-El (left) and Range-Az (right) for $C_B = 0.7$

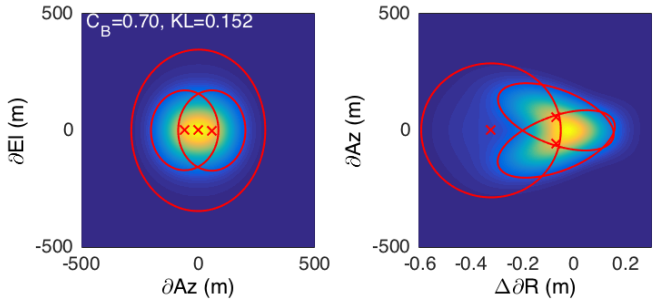

Figure 61: GM Marginal PDFs with $N_G = 3$ 2R-T for Az-El (left) and Range-Az (right) for $C_B = 0.7$

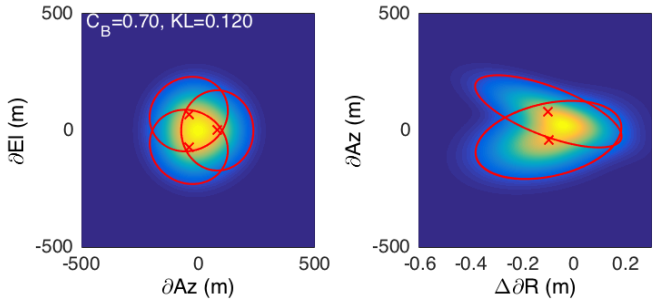

Figure 62: GM Marginal PDFs with $N_G = 3$ 3R for Az-El (left) and Range-Az (right) for $C_B = 0.7$

4.3.3 $N_G = 3$ Solutions

For the $N_G = 3$ case, the same measurement with bias significance of $C_B = 0.7$ was fitted. Again, two main geometries are present. Adding a range tail to 2R results in 2R-T shown in Figure 61 and adding another component to the ring results in 3R shown in Figure 60. In this case, 3R has 25% better KL divergence performance at 0.120, compared to 2R-T with 0.152, a reversal of the situation for $N_G = 2$.

4.3.4 $N_G = 4$ Solutions

For $C_B = 0.7$ and $N_G = 4$ there are three regular geometries. Adding a range tail to 3R gives 3R-T in Figure 63, with 0.086 KL divergence, and adding a central component gives C-3R in Figure 64, with 0.087 KL divergence (nearly identical). However, the optimal geometry for this bias significance is 4R in Figure 65, with 0.077 KL divergence.

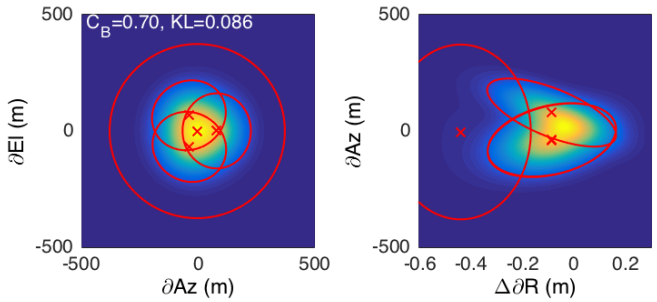

Figure 63: GM Marginal PDFs with $N_G = 4$ 3R-T for Az-El (left) and Range-Az (right) for $C_B = 0.7$

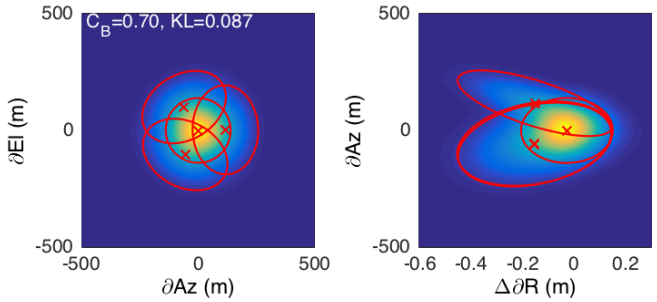

Figure 64: GM Marginal PDFs with $N_G = 4$ C-3R for Az-El (left) and Range-Az (right) for $C_B = 0.7$

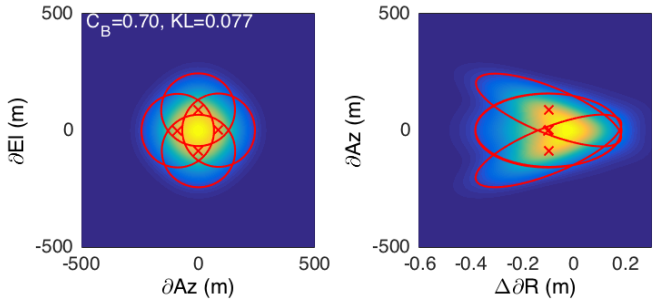

Figure 65: GM Marginal PDFs with $N_G = 4$ 4R for Az-El (left) and Range-Az (right) for $C_B = 0.7$

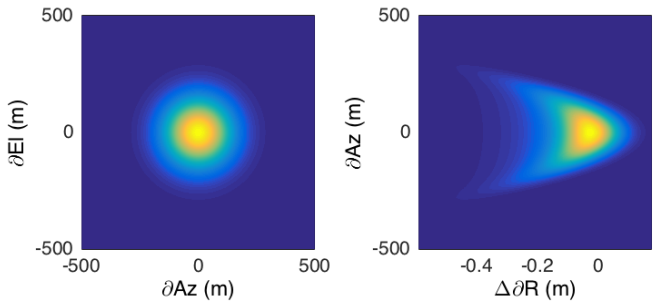

Figure 66: True Marginal PDFs for Az-El (left) and Range-Az (right) for $C_B = 1.2$

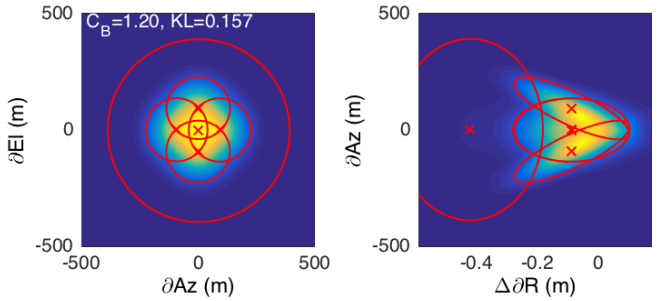

Figure 67: GM Marginal PDFs with $N_G = 5$ 4R-T for Az-El (left) and Range-Az (right) for $C_B = 1.2$

4.3.5 $N_G = 5$ Solutions

As $N_G = 4$ gives a fairly low KL divergence for $C_B = 0.7$, the measurement bias significance was increased to $C_B = 1.2$ for $N_G = 5$. The true marginals for this distribution are shown in Figure 66. The pattern of solution evolution from $N_G = 4$ to $N_G = 5$ is similar to that from $N_G = 3$ to $N_G = 4$. Adding a range tail to 4R gives 4R-T in Figure 67, with KL divergence 0.157, and adding a range tail to C-3R gives C-3R-T in Figure 68, with KL divergence 0.184. However, the optimal geometry comes not from adding a range tail, but adding another component to the ring in C-3R to get C-4R in Figure 69, with KL divergence 0.149. Furthermore, although 5R seems like a geometry worth exploring, this does not represent a stable local optimum for EM - the algorithm converges to C-4R in the case of a 5R initialization.

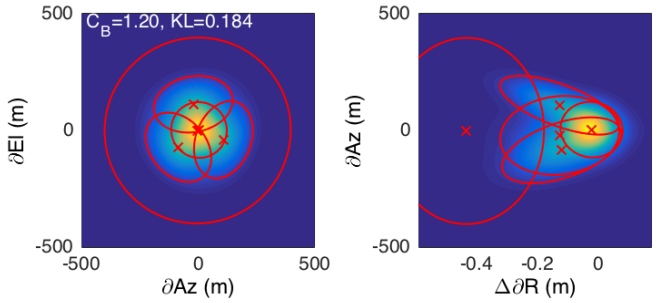

Figure 68: GM Marginal PDFs with $N_G = 5$ C-3R-T for Az-El (left) and Range-Az (right) for $C_B = 1.2$

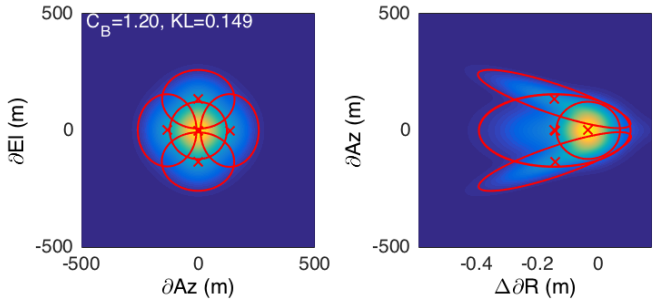

Figure 69: GM Marginal PDFs with $N_G = 5$ C-4R for Az-El (left) and Range-Az (right) for $C_B = 1.2$

4.3.6 $N_G = 6$ Solutions

For $N_G = 6$, a bias significance of $C_B = 1.2$ was retained. Although 5R was not found to be a stable local optimum for $N_G = 5$, for $N_G = 6$ 5R-T is a stable local optimum, shown in Figure 70. However, this solution is inferior in KL divergence (0.129) to the alternative range tail solution C-4R-T shown in Figure 71 (0.118). This is somewhat expected, given that 5R converges to C4R for $N_G = 5$. Ultimately, the C-5R solution with no range tail is optimal with KL divergence 0.108. As in the $N_G = 5$ case, the simple ring 6R is not a stable local optimum, and this initialization converges to C-5R.

4.3.7 $N_G = 7$ Solutions

A bias significance of $C_B = 1.2$ was also used for $N_G = 7$. Following the intuition of the results in $N_G = 6$, 6R-T was tested and found to be a stable local solution,

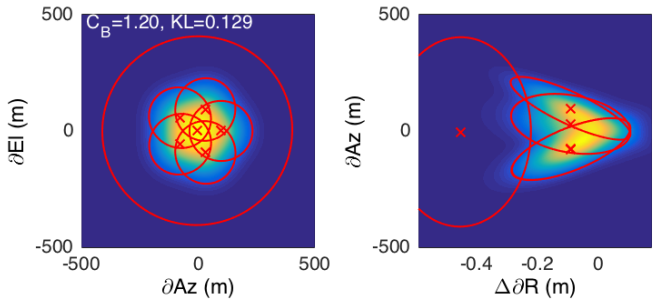

Figure 70: GM Marginal PDFs with $N_G = 6$ 5R-T for Az-El (left) and Range-Az (right) for $C_B = 1.2$

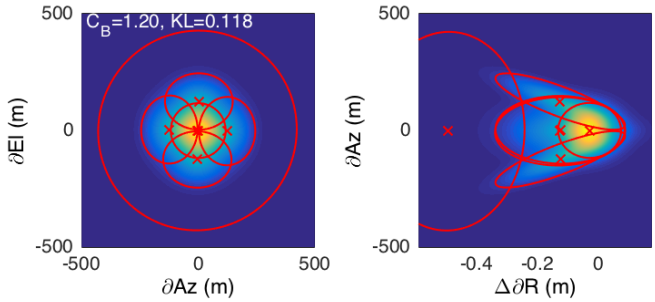

Figure 71: GM Marginal PDFs with $N_G = 6$ C-4R-T for Az-El (left) and Range-Az (right) for $C_B = 1.2$

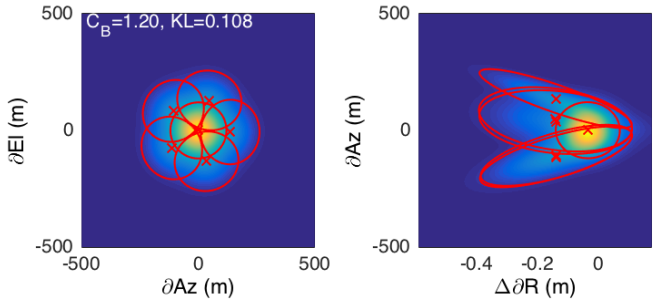

Figure 72: GM Marginal PDFs with $N_G = 6$ C-5R for Az-El (left) and Range-Az (right) for $C_B = 1.2$

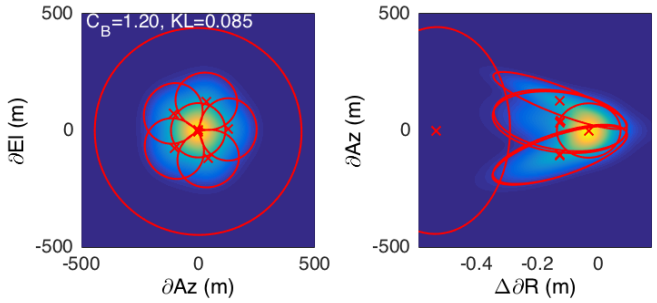

Figure 73: GM Marginal PDFs with $N_G = 7$ C-5R-T for Az-El (left) and Range-Az (right) for $C_B = 1.2$

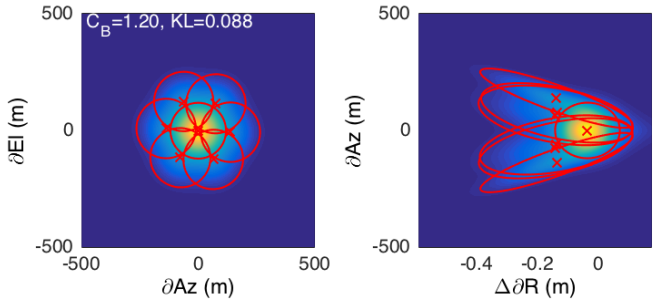

Figure 74: GM Marginal PDFs with $N_G = 7$ C-6R for Az-El (left) and Range-Az (right) for $C_B = 1.2$

however it is inferior to C-5R-T in Figure 73. The stable solution without a range tail is C-6R in Figure 74. Unlike in the $N_G = 6$ case, these solutions have a very similar KL divergence of 0.085 and 0.088 respectively, which is almost within sampling error. Both solutions are only a slight improvement in KL divergence on their corresponding $N_G = 6$ solutions, indicating that further extension of these geometric pattern may provide only marginal benefit.

4.3.8 $N_G = 8$ Solutions

For $N_G = 8$, the bias significance was again increased to $C_B = 2$. The marginals for this measurement are shown in Figure 75. Adding a range tail to C-6R gives C-6R-T in Figure 76 with KL divergence 0.169. However, continuing the pattern for the non-range-tail case by augmenting C-6R to C-7R in Figure 77 gives a highly sub-optimal solution with KL divergence 0.211. Through random search, an alternative solution

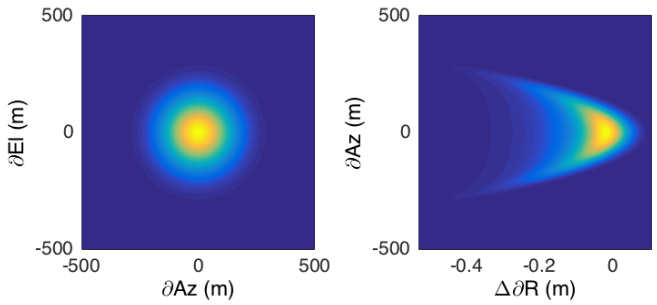

Figure 75: True Marginal PDFs for Az-El (left) and Range-Az (right) for $C_B = 2.0$

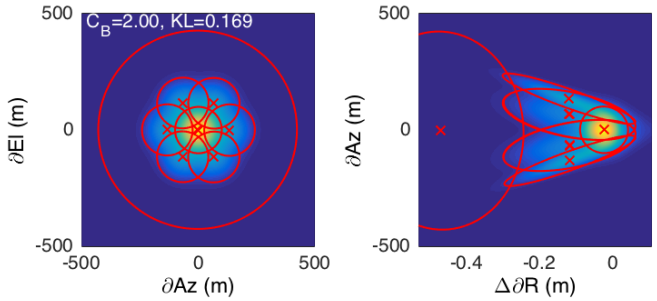

Figure 76: GM Marginal PDFs with $N_G = 8$ C-6R-T for Az-El (left) and Range-Az (right) for $C_B = 2.0$

with no range tail 3R-5R can be found. This solution is shown in Figure 78 and has only slightly suboptimal KL divergence 0.180. However, the 5 components in the outer ring are arranged in irregular intervals, so that this does not present a desirable solution from an ease-of-modeling perspective. By extending this geometry to 3R-6R, a more regular solution may be possible.

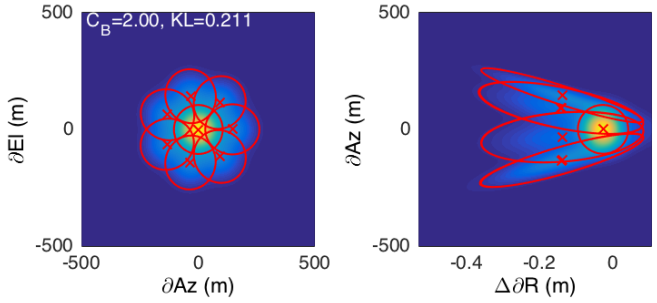

Figure 77: GM Marginal PDFs with $N_G = 8$ C-7R for Az-El (left) and Range-Az (right) for $C_B = 2.0$

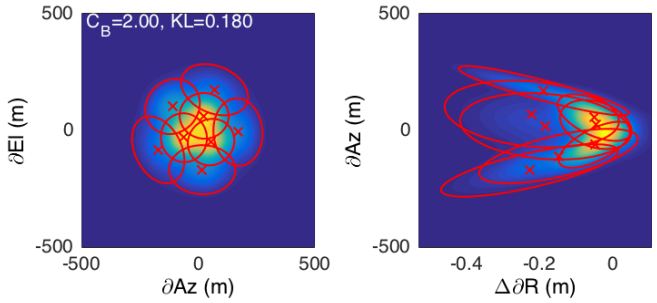

Figure 78: GM Marginal PDFs with $N_G = 8$ 3R-5R for Az-El (left) and Range-Az (right) for $C_B = 2.0$

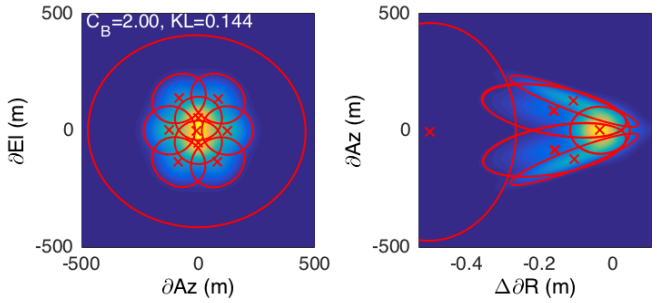

Figure 79: GM Marginal PDFs with $N_G = 9$ 2R-6R-T for Az-El (left) and Range-Az (right) for $C_B = 2.0$

4.3.9 $N_G = 9$ Solutions

For $N_G = 9$, the bias significance was kept at $C_B = 2$. As expected, from the $N_G = 8$ results, 3R-6R forms a stable, regular geometry with no range tail shown in Figure 81. A KL divergence of 0.145 is achieved. For the range tail cases, C-7R-T in Figure 80 has a slightly inferior KL divergence of 0.149, whereas removing one of the central components of 3R-6R to get 2R-6R-T in Figure 79 has a comparable KL divergence 0.144 (within sampling error). This serves to further illustrate that C-7R is a suboptimal base geometry, as was seen in the $N_G = 8$ case. Although the outer ring of 3R-6R does not contain 6 isomorphic components, it has two groups of 3 isomorphic components, which is sufficiently regular to be stored in a modeling database.

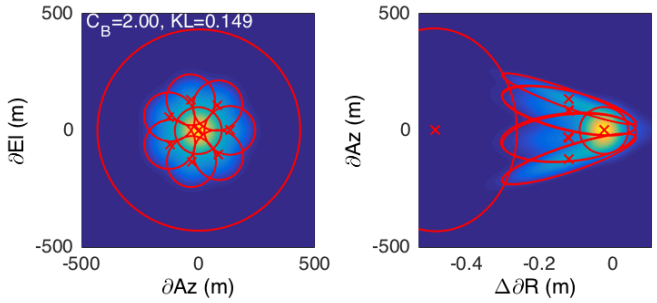

Figure 80: GM Marginal PDFs with $N_G = 9$ C-7R-T for Az-El (left) and Range-Az (right) for $C_B = 2.0$

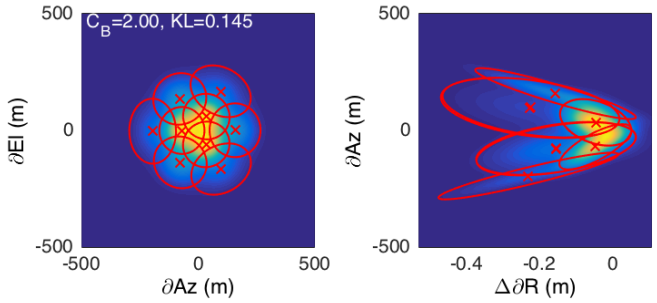

Figure 81: GM Marginal PDFs with $N_G = 9$ 3R-6R for Az-El (left) and Range-Az (right) for $C_B = 2.0$

4.3.10 $N_G = 10$ and Higher Solutions

Based on the previous results, it seems clear that the useful geometries at higher N_G must consist of more than a single ring optionally arranged around an inner component and/or range tail. The question is how many components should exist in each ring, and whether the resulting geometries are sufficiently regular. While components in an inner ring arranged around a center component tend to be isomorphic, the same may not be true of components in an outer ring arranged around this.

Again, a measurement with $C_B = 2$ was used to study $N_G = 10$ through $N_G = 12$. For $N_G = 10$, based on previous results, two solutions without a range tail seem to be logical, C-3R-6R and 4R-6R. It was found that 4R-6R is the stable solution, as a C-3R-6R initialization converges to this. The 4R-6R solution is shown in Figure 82 and has a KL divergence of 0.125, which is comparable to the 0.119 achieved by the range tail

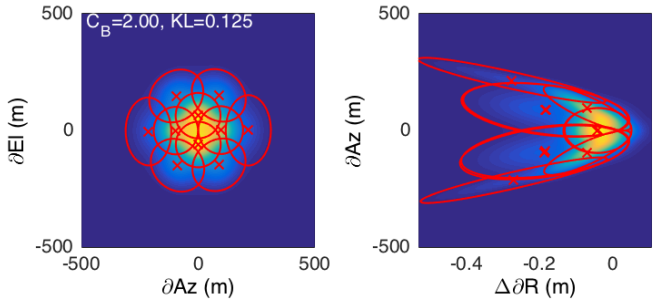

Figure 82: GM Marginal PDFs with $N_G = 10$ 4R-6R for Az-El (left) and Range-Az (right) for $C_B = 2.0$

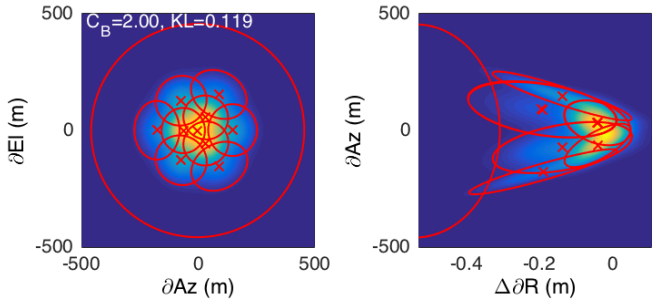

Figure 83: GM Marginal PDFs with $N_G = 10$ 3R-6R-T for Az-El (left) and Range-Az (right) for $C_B = 2.0$

solution 3R-6R-T shown in Figure 83. The solution 5R-5R was also considered and is shown in Figure 84. However, this solution has sub-optimal performance compared to 4R-6R, and the 5 inner components do not cluster in a regular manner due to the lack of a central component to enforce isomorphism.

For $N_G = 11$, promoting 4R-6R to a range tail solution 4R-6R-T in Figure 85 has a KL divergence of 0.106. For a non-range-tail solution, there are two possible regular options, C-5R-5R and C-4R-6R. Both of these yield stable solutions shown in Figure 86 and 87, but C-5R-5R has slightly better KL divergence at 0.105 as well as being a more regular solution (isomorphic components in both rings).

The obvious range tail solution for $N_G = 12$ is C-5R-5R-T with 0.096 KL divergence, as shown in Figure 88. However, the no tail solution 4R-8R appears to slightly outperform this with 0.093 KL divergence, as shown in Figure 89. The 4R-8R solution is isomorphic in the inner ring and has two sets of isomorphic components in the

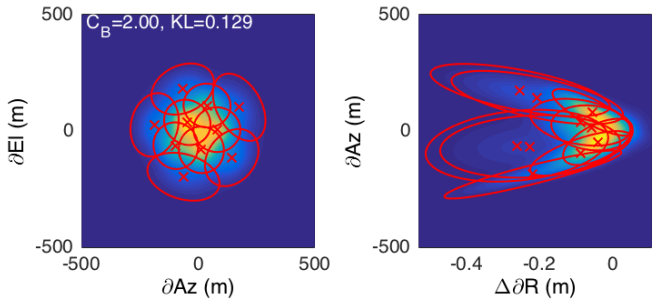

Figure 84: GM Marginal PDFs with $N_G = 10$ 5R-5R for Az-El (left) and Range-Az (right) for $C_B = 2.0$

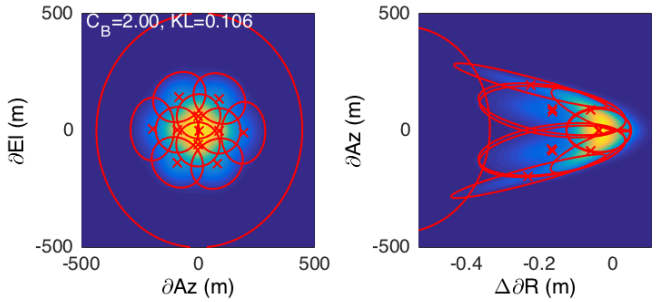

Figure 85: GM Marginal PDFs with $N_G = 11$ 4R-6R-T for Az-El (left) and Range-Az (right) for $C_B = 2.0$

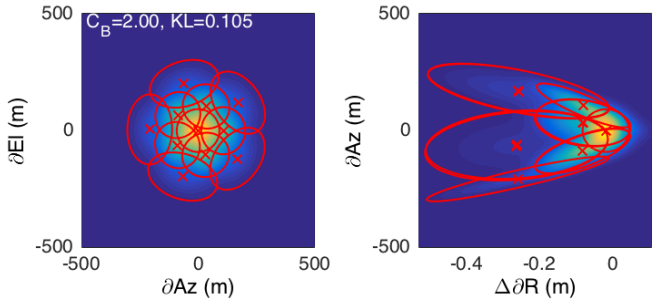

Figure 86: GM Marginal PDFs with $N_G = 11$ C-5R-5R for Az-El (left) and Range-Az (right) for $C_B = 2.0$

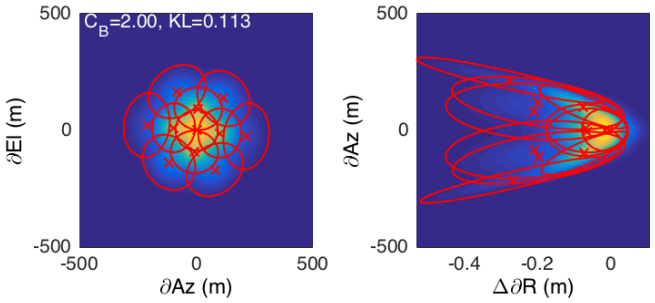

Figure 87: GM Marginal PDFs with $N_G = 11$ C-4R-6R for Az-El (left) and Range-Az (right) for $C_B = 2.0$

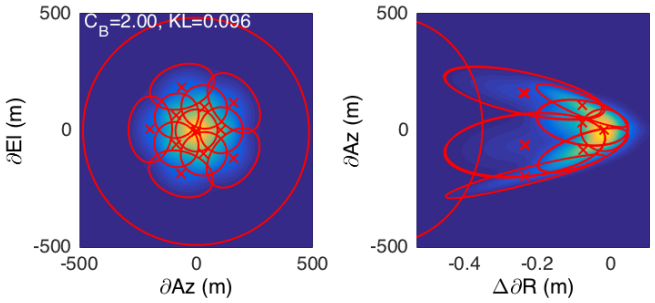

Figure 88: GM Marginal PDFs with $N_G = 12$ C-5R-5R-T for Az-El (left) and Range-Az (right) for $C_B = 2.0$

outer ring. Therefore it is a reasonable solution from a regularity perpective as well.

For $N_G = 13$, the bias significance of the measurement was increased to $C_B = 3$. The true marginal distributions of this measurement are shown in Figure 90. Two regular no-tail geometries can be proposed for $N_G = 13$, C-R4-R8 and C-R6-R6. It was found that C-R4-R8 is not a stable geometry, and initialization with this geometry converges to C-R5-R7 in Figure 92, which is irregular. C-R6-R6 in Figure 93 does converge and has superior KL divergence performance to C-5R-7R (0.155 vs 0.162) and two rings of isomorphic components.

4.3.11 Preferred Geometries for Measurement Modeling

The results above show that multiple stable solution geometries exist for GM fits of a given size. In general, these solutions may be categorized into two major classes,

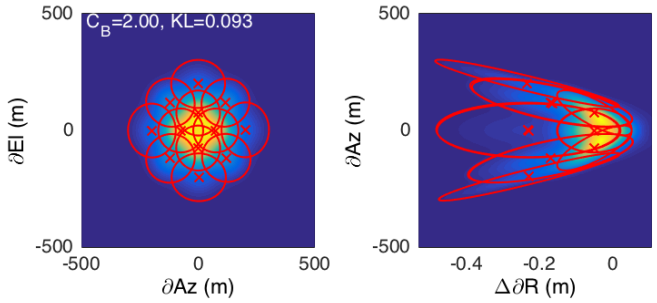

Figure 89: GM Marginal PDFs with $N_G = 12$ 4R-8R for Az-El (left) and Range-Az (right) for $C_B = 2.0$

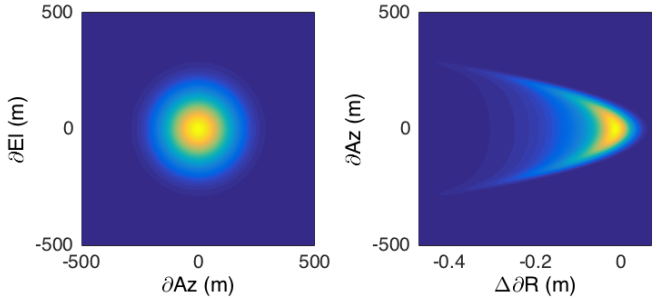

Figure 90: True Marginal PDFs for Az-El (left) and Range-Az (right) for $C_B = 3.0$

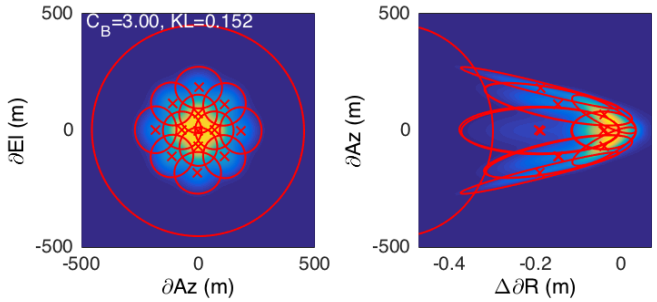

Figure 91: GM Marginal PDFs with $N_G = 13$ 4R-8R-T for Az-El (left) and Range-Az (right) for $C_B = 3.0$

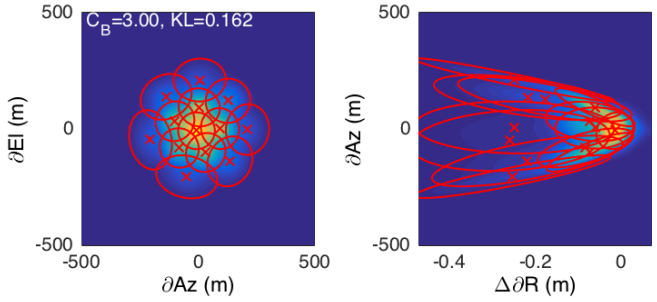

Figure 92: GM Marginal PDFs with $N_G = 13$ C-5R-7R for Az-El (left) and Range-Az (right) for $C_B = 3.0$

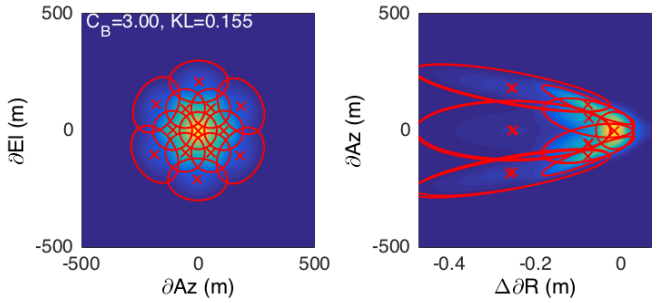

Figure 93: GM Marginal PDFs with $N_G = 13$ C-6R-6R for Az-El (left) and Range-Az (right) for $C_B = 3.0$

those having a range tail component and those without. For use in a Gaussian Mixture Kalman filter the solutions without a range tail component are preferable, since the cross-range support region of each individual component is smaller. Therefore, the cross-range support of the state prior distribution component in the next filter iteration will reduced, allowing the measurement pre-conditioning method described in Section 2.4.4 to choose a smaller sized representation for each conditioned Cartesian measurement.

Furthermore, it is seen that some numbers of components do not yield a solution with sufficient geometric regularity (and no range tail) to be efficiently represented in a lookup table (for example $N_G = 14$). Table 5 shows the the preferred geometry associated with numbers of components up to $N_G = 20$. If a given number of components does not admit an efficient regular geometry, it is left blank.

4.4 Parameterization of Gaussian Mixture Geometries with Bias Significance Ratios of One

As in [9] is is desired to build a lookup table of normalized parameters so that a Gaussian mixture of a given size N_G may be rapidly generated to model a measurement with given spherical mean μ_S and diagonal covariance Σ_S. In addition, the KL divergence of the fit is stored in the table and used to predict the performance of the model. Based on the results of Section 4.2, it was seen that measurements with

Table 5: Preferred Mixture Geometry Solutions for 3D Measurements

N_G	Geometry Code	N_G	Geometry Code
		11	C-5R-5R
2	2R	12	4R-8R
3	3R	13	C-6R-6R
4	4R	14	
5	C-4R	15	C-7R-7R
6	C-5R	16	C-6R-9R
7	C-6R	17	
8	C-7R	18	
9	3R-6R	19	C-6R-9R-3R
10	4R-6R	20	4R-8R-8R

equal C_B and R_B have equal KL divergence performance when fitted using the EM algorithm. Additionally, the weights of the components in the fit, ω_k, are the same. Therefore, as in the 2-D case of [9] it should be possible to transform the means and covariances, μ_k and Σ_k, into a normalized form which is constant for a given C_B, R_B, and geometric solution having N_G components. This study is restricted to the circularly symmetric case of $R_B = 1$.

In Section 4.3, it was seen that many of the preferred geometries of mixture components include isomorphic classes of components distributed regularly in a circular "ring" (when viewed in the Az/El plane). Therefore, the following set of transformations is proposed to efficiently represent this geometric structure.

Let $\{\omega_k, \mu_k, \Sigma_k\}_{k=1}^{N_G}$ represent the weights, means, and covariances of the measurement mixture components in Cartesian space. First, in a similar way to that in Section 2.3.4, transform the means and covariances to spherical coordinates using the Jacobian of the transformation linearized around the mean. Thus

$$r_k = |\mu_k| \tag{228}$$

$$\alpha_k = \tan^{-1} \frac{[\mu_k]_x}{[\mu_k]_y} \tag{229}$$

$$\varepsilon_k = \sin^{-1} \frac{[\mu_k]_z}{|\mu_k|} \tag{230}$$

$$\Sigma_{S,k} = J_{S,k}^{-1} \Sigma_k J_{S,k}^{-T} \tag{231}$$

where the notation $[\mu_k]_{(\cdot)}$ indicates a particular vector component of μ_k, and $J_{S,k}$ is equal to the Jacobian (191) evaluated at $\begin{bmatrix} r_k & \alpha_k & \varepsilon_k \end{bmatrix}^T$. Next, calculate the displacement of the spherical component means from the original distribution mean.

$$\Delta r_k = r_k - \bar{r} \qquad \Delta \alpha_k = \alpha_k - \bar{\alpha} \qquad \Delta \varepsilon_k = \varepsilon_k - \bar{\varepsilon} \tag{232}$$

During this step, perform any unwrapping of angles to minimize the absolute value of the differences $\Delta\alpha_k$ and $\Delta\varepsilon_k$. The spherical coordinate differences are now further normalized by the standard deviations of the original distribution according to

$$\Delta \hat{r}_k = \frac{\Delta r_k}{\sqrt{2}\sigma_r C_B} \qquad \Delta \hat{\alpha}_k = \frac{\Delta \alpha_k}{\sigma_\alpha} \qquad \Delta \hat{\varepsilon}_k = \frac{\Delta \varepsilon_k}{\sigma_\varepsilon} \tag{233}$$

where $\sqrt{2}\sigma_r C_B$ is the expected range shift in the mean from (227) for a single Gaussian moment match to a measurement with $R_B = 1$. Therefore, in the single Gaussian case these parameters are expected to be $\Delta\hat{r}_k = -1$, $\Delta\hat{\alpha}_k = 0$, $\Delta\hat{\varepsilon}_k = 0$. At this point, $\Delta\hat{r}_k$ for all components in an isomorphic ring class are approximately equal. Furthermore, for any central component, $\Delta\hat{\alpha}_k$ and $\Delta\hat{\varepsilon}_k$ are approximately zero. Since the range tail component is not used in any of the preferred geometries, it remains to parameterize the angular positions of the isomorphic rings. Considering the normalized Az-El plane as a pseudo-Cartesian system, define the Cartesian-to-polar transformation

$$\Delta \hat{\theta}_k = \sqrt{\Delta \hat{\alpha}_k^2 + \Delta \hat{\varepsilon}_k^2} \tag{234}$$

$$\psi_k = \tan^{-1} \frac{\Delta \hat{\alpha}_k}{\Delta \hat{\varepsilon}_k} \tag{235}$$

where $\Delta\hat{\theta}_k$ is the "angular displacement" and ψ_k is the "angular orientation." The Jacobian of this transformation is

$$J_{\theta\psi,k} = \frac{1}{\Delta \hat{\theta}_k} \begin{bmatrix} \Delta \hat{\alpha}_k & \frac{\Delta \hat{\varepsilon}_k}{\Delta \hat{\theta}_k} \\ \Delta \hat{\varepsilon}_k & -\frac{\Delta \hat{\alpha}_k}{\Delta \hat{\theta}_k} \end{bmatrix} \tag{236}$$

Now, $\Delta\hat{\theta}_k$ is equal for all components within an isomorphic ring class, whereas ψ_k is spaced at regular intervals. If the total bias significance C_B changes, $\Delta\hat{\theta}_k$ will change, but the ψ_k spacing will remain the same.

For the covariance parameterization, define the scaling matrix

$$S = \begin{bmatrix} \sigma_r\sqrt{1+2C_B^2} & 0 & 0 \\ 0 & \sigma_\alpha & 0 \\ 0 & 0 & \sigma_\varepsilon \end{bmatrix} \tag{237}$$

and compute the normalized covariance matrix according to

$$\hat{\Sigma}_{S,k} = S^{-1}\Sigma_{S,k}S^{-1} \tag{238}$$

Then the parameters of the normalized covariance can be expressed as

$$\begin{aligned} \hat{\sigma}_{r,k} &= \sqrt{[\hat{\Sigma}_{S,k}]_{rr}} & \rho_{r\alpha,k} &= \frac{[\hat{\Sigma}_{S,k}]_{r\alpha}}{\hat{\sigma}_{r,k}\hat{\sigma}_{\alpha,k}} \\ \hat{\sigma}_{\alpha,k} &= \sqrt{[\hat{\Sigma}_{S,k}]_{\alpha\alpha}} & \rho_{r\varepsilon,k} &= \frac{[\hat{\Sigma}_{S,k}]_{r\varepsilon}}{\hat{\sigma}_{r,k}\hat{\sigma}_{\varepsilon,k}} \\ \hat{\sigma}_{\varepsilon,k} &= \sqrt{[\hat{\Sigma}_{S,k}]_{\varepsilon\varepsilon}} & \rho_{\alpha\varepsilon,k} &= \frac{[\hat{\Sigma}_{S,k}]_{\alpha\varepsilon}}{\hat{\sigma}_{\alpha,k}\hat{\sigma}_{\varepsilon,k}} \end{aligned} \tag{239}$$

where the range standard deviation scaling $\sigma_r\sqrt{1+2C_B^2}$ is equivalent to the range standard deviation for a single-Gaussian moment matched distribution so that in the single-Gaussian case $\hat{\sigma}_{r,k} = \hat{\sigma}_{\alpha,k} = \hat{\sigma}_{\varepsilon,k} = 1$. For any central component, the correlation coefficients are expected to be statistically equal to zero and $\hat{\sigma}_{\alpha,k} = \hat{\sigma}_{\varepsilon,k}$ due to circular symmetry, so these parameters may be represented in the table as a unified parameter $\hat{\sigma}_{\alpha\varepsilon,k}$.

Therefore, the parameters ω_k, $\Delta\hat{r}_k$, $\hat{\sigma}_{r,k}$, and $\hat{\sigma}_{\alpha,k} = \hat{\sigma}_{\varepsilon,k} = \hat{\sigma}_{\alpha\varepsilon,k}$ may be used to represent a central component in a normalized parameter lookup table for 3D components, where the parameters are stored for a given N_G, geometry arrangement, C_B, and $R_B = 1$.

For the ring components, further transform the covariance $\hat{\Sigma}_{S,k}$ into the angle displacement-orientation frame centered on each mean through multiplication by the

Jacobian according to

$$\hat{\Sigma}_{\theta\psi,k} = \begin{bmatrix} 1 & 0_{1\times 2} \\ 0_{2\times 1} & J_{\theta\psi,k} \end{bmatrix} \hat{\Sigma}_{S,k} \begin{bmatrix} 1 & 0_{1\times 2} \\ 0_{2\times 1} & J_{\theta\psi,k}^T \end{bmatrix} \tag{240}$$

and define

$$\begin{aligned} \hat{\sigma}_{r,k} &= \sqrt{[\hat{\Sigma}_{\theta\psi,k}]_{rr}} & \rho_{r\hat{\theta},k} &= \frac{[\hat{\Sigma}_{\theta\psi,k}]_{r\hat{\theta}}}{\hat{\sigma}_{r,k}\hat{\sigma}_{\hat{\theta},k}} \\ \hat{\sigma}_{\hat{\theta},k} &= \sqrt{[\hat{\Sigma}_{\theta\psi,k}]_{\hat{\theta}\hat{\theta}}} & \rho_{r\psi,k} &= \frac{[\hat{\Sigma}_{\theta\psi,k}]_{r\psi}}{\hat{\sigma}_{r,k}\hat{\sigma}_{\psi,k}} \\ \hat{\sigma}_{\psi,k} &= \sqrt{[\hat{\Sigma}_{\theta\psi,k}]_{\psi\psi}} & \rho_{\hat{\theta}\psi,k} &= \frac{[\hat{\Sigma}_{\theta\psi,k}]_{\hat{\theta}\psi}}{\hat{\sigma}_{\hat{\theta},k}\hat{\sigma}_{\psi,k}} \end{aligned} \tag{241}$$

These are already normalized parameters since the definitions of $\hat{\theta}$ and ψ are based on the normalized differences in Az and El.

Due to the circular symmetry of the ring, it is expected that $\rho_{r\psi,k} = \rho_{\hat{\theta}\psi,k} = 0$. Furthermore, all of the parameters except ψ_k are expected to be statistically equivalent for all isomorphic components in the ring. If I_ℓ is the set of component indices in ring ℓ, then one set of parameters ω_ℓ, $\Delta\hat{r}_\ell$, $\Delta\hat{\theta}_\ell$, $\hat{\sigma}_{r,\ell}$, $\hat{\sigma}_{\theta,\ell}$, $\hat{\sigma}_{\psi,\ell}$, and $\rho_{r\theta,\ell}$ may be used to represent all the components in a ring with only ψ_k varying over the indices in the ring set. This greatly reduces the size of the representation in the table.

4.4.1 Examples of Transformation Results

In order to illustrate the effectiveness of this sequence of transformations in representing GM solutions for $R_B = 1$, two cases with an $N_G = 6$ mixture solution with geometry C-5R are fitted to a measurement with $\bar{r} = 100$ km, $C_B = 1.5$, and $R_B = 1$. In the first case, $\sigma_\varepsilon = 1$ mrad and $\varepsilon = 0.4\pi$ rad resulting in $\sigma_\alpha = 5.12$ mrad and $\sigma_r = 4.714$ cm, whereas in the second case $\sigma_\varepsilon = 10$ mrad and $\varepsilon = 0$ rad resulting in $\sigma_\alpha = 10$ mrad and $\sigma_r = 4.714$ m.

For the first case, the Cartesian projection of the mixture components is shown in Figure 94, where ∂R, ∂Az, ∂El are the Jacobian axes defined in Section 4.3 and

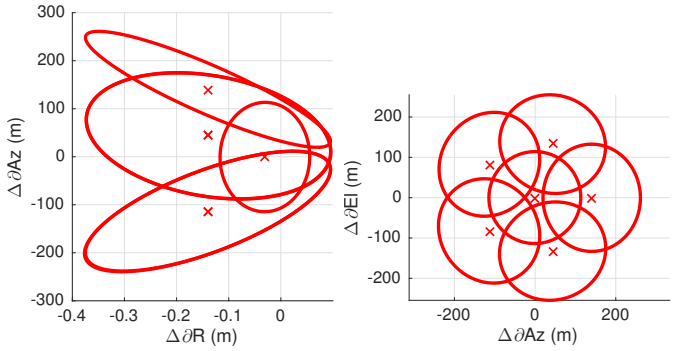

Figure 94: Cartesian Component Layout for C-5R with $C_B = 1.5$, $\sigma_\varepsilon = 10^{-3}$ rad, and $\varepsilon = 0.4\pi$ rad

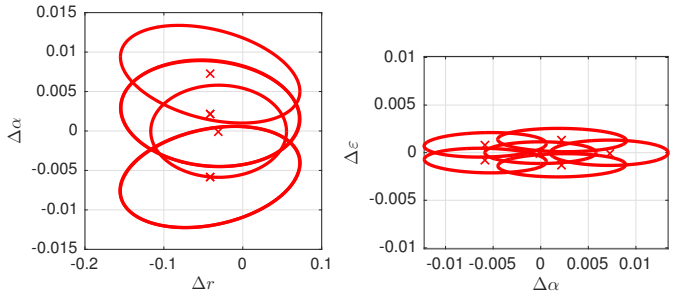

Figure 95: Spherical Component Layout for C-5R with $C_B = 1.5$, $\sigma_\varepsilon = 10^{-3}$ rad, and $\varepsilon = 0.4\pi$ rad

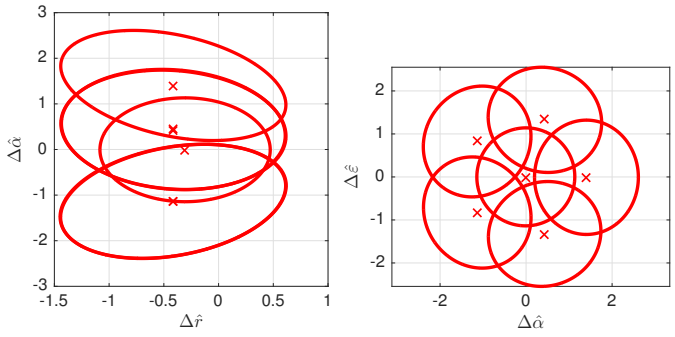

Figure 96: Normalized Spherical Component Layout for C-5R with $C_B = 1.5$, $\sigma_\varepsilon = 10^{-3}$ rad, and $\varepsilon = 0.4\pi$ rad

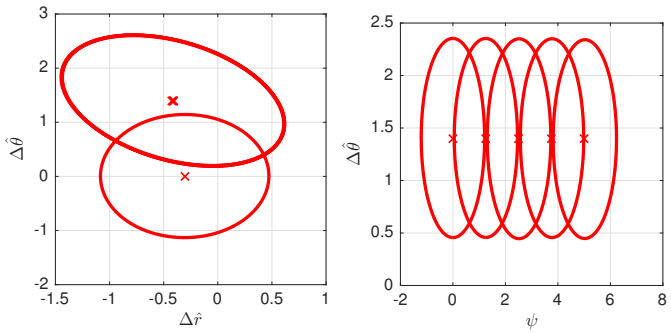

Figure 97: Rotated Spherical Component Layout for C-5R with $C_B = 1.5$, $\sigma_\varepsilon = 10^{-3}$ rad, and $\varepsilon = 0.4\pi$ rad

the Δ prefix indicates difference from the spherical to Cartesian conversion of the spherical measurement mean. The marginal means are indicated by a cross, while the covariance ellipses show the 2-sigma (Mahalanobis distance of 4) contours of the marginal distribution in the indicated plane. Even though $\sigma_\alpha \neq \sigma_\epsilon$ in this case due to the $\cos\varepsilon$ factor applied to the ∂Az axis in Cartesian space, the components have circular symmetry in the ∂Az-∂El plane.

Figure 95 shows the converted spherical mean displacements Δr_k, $\Delta\alpha_k$, and $\Delta\varepsilon_k$. as well as the contours of the spherical covariances $\Sigma_{S,k}$ as covariance ellipses. In this coordinate system the mismatch in azimuth and elevation error is clearly visible as the covariance ellipses appear extremely stretched in the Az-El plane. However, once normalization is applied to yield $\Delta\hat{r}_k$, $\Delta\hat{\alpha}_k$, $\Delta\hat{\varepsilon}_k$, and $\hat{\Sigma}_{S,k}$, the circular symmetry of the geometry returns as shown in Figure 96. Finally, when the components are converted to angle displacement and orientation ($\hat{\theta}_k$,ψ_k, and $\hat{\Sigma}_{\theta\psi,k}$) as shown in Figure 97, it becomes evident that all the components in the ring have the same parameters (within sampling error) except for ψ_k. (The central component is included only in the left plot of the figure but not in the right since the Jacobian axis for ψ_k is singular at $\Delta\hat{\theta}_k = 0$.)

Likewise, Figure 98 shows the Cartesian projection of a C-5R fit to the second case ($\sigma_\varepsilon = 10$ mrad and $\varepsilon = 0$ rad). By comparison to the first case in Figure 94, the

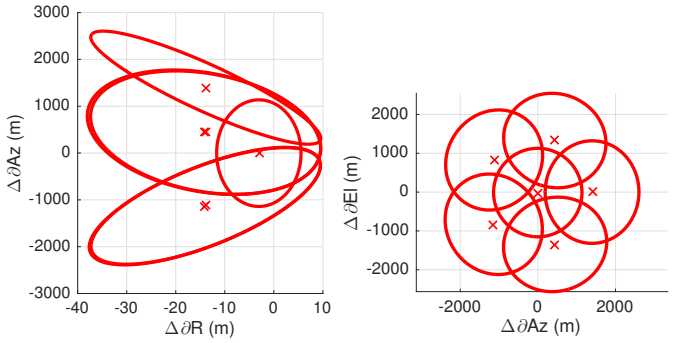

Figure 98: Cartesian Component Layout for C-5R with $C_B = 1.5$, $\sigma_\varepsilon = 10^{-2}$ rad, and $\varepsilon = 0\pi$ rad

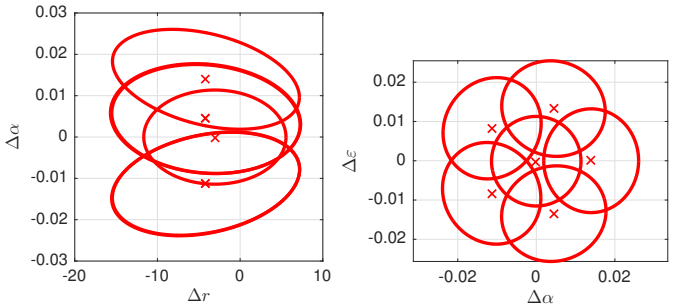

Figure 99: Spherical Component Layout for C-5R with $C_B = 1.5$, $\sigma_\varepsilon = 10^{-2}$ rad, and $\varepsilon = 0\pi$ rad

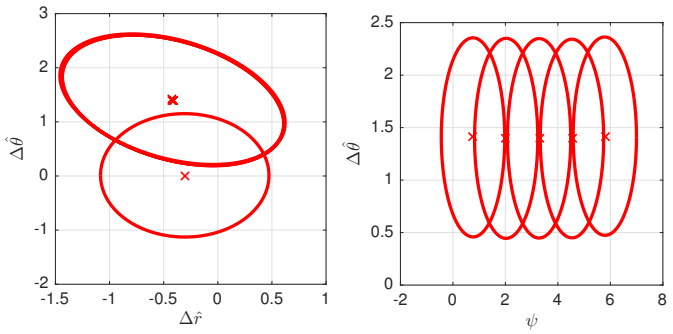

Figure 100: Rotated Spherical Component Layout for C-5R with $C_B = 1.5$, $\sigma_\varepsilon = 10^{-2}$ rad, and $\varepsilon = 0\pi$ rad

geometric arrangement of the components is nearly identical with different scaling of the axes. When converted to Δr_k, $\Delta\alpha_k$, and $\Delta\varepsilon_k$ spherical coordinates in Figure 99 the containment regions for $\Delta\alpha_k$ and $\Delta\varepsilon_k$ have circular symmetry unlike in the case of Figure 95. This is due to the fact that the elevation angle is zero, meaning $\sigma_\alpha = \sigma_\varepsilon$, whereas in the first case $\sigma_\alpha > \sigma_\varepsilon$ to account for the cosine term in the azimuth bias significance. For brevity, the normalized spherical coordinate plot is not shown for the second case, but the final result for the normalized rotated coordinates is shown in Figure 100. By comparison to Figure 97, the result is almost identical, except for a slight shift in angle ϕ_k which causes the first ring component to wrap on the opposite side of 2π. However, due to the circular symmetry of the solution at $R_B = 1$, it is expected that a shift in the overall angle orientation ϕ_k will have little to no effect on the KL divergence of the solution, therefore the results of the EM algorithm applied to this problem may show various shifts in ϕ_k. In this case, the differences between ϕ_k should remain consistent, as they do.

Overall, the result of this demonstration is to verify that the transformed solution parameters for these two measurements with the same C_B and R_B but different σ_r, σ_α, σ_ε and $\bar{\varepsilon}$ are statistically identical. This indicates that the transformed parameter set for these mixture solutions may be used to build an efficient lookup table based only on C_B, R_B, and a selected geometry with N_G components.

Tables 6 through 8 show the parameters that result from this fitting example. In these tables, the index expression $(\cdot)_c$ indicates the parameters associated with the central component, while $(\cdot)_\ell$ indicates the parameters associated with ring ℓ. (In this C-5R $N_G = 6$ example, there is only one ring, so only $(\cdot)_1$ is used.) Table 6 shows the KL divergence achieved by the C-5R solution for various levels of total bias significance. Additionally, it shows the weight, mean, and covariance parameters associated with the central component. The central component parameters $\Delta\hat{\alpha}_c$ and $\Delta\hat{\varepsilon}_c$ are zero as described in the previous analysis. Also, all cross-correlation values

for the central component are defined to be zero, and the normalized azimuth and elevation variances are identical and expressed by a single parameter $\hat{\sigma}_{\alpha\varepsilon,c}$.

Table 7 shows the weight, mean, and covariance parameters associated with the components in Ring 1. For the ring components, only the range-to-angle displacement correlation is non-zero, and is given by $\rho_{r\hat{\theta},\ell}$. Additionally, Table 8 shows the angular orientation values for Ring 1 which are associated with the mean positions. As expected for a simple ring structure with five components, the orientation values are equally spaced around a circle at $\frac{2\pi}{5}$ intervals. The angle orientation values do not change with bias significance.

Table 6: KL Divergence and Center Parameters for C-5R (sample)

C_B	KL	ω_c	$\Delta\hat{r}_c$	$\hat{\sigma}_{r,c}$	$\hat{\sigma}_{\alpha\varepsilon,c}$
0.60	0.021	0.380	−0.45	0.70	0.69
0.70	0.031	0.378	−0.42	0.64	0.68
0.80	0.044	0.375	−0.40	0.59	0.66
0.90	0.058	0.373	−0.38	0.55	0.64
1.00	0.074	0.370	−0.36	0.51	0.63
1.10	0.091	0.368	−0.35	0.48	0.61
1.20	0.109	0.367	−0.34	0.45	0.60

The results show that the central component is heavily weighted compared to the

Table 7: Ring 1 Parameters for C-5R (sample)

C_B	ω_1	$\Delta\hat{r}_1$	$\Delta\hat{\theta}_1$	$\hat{\sigma}_{r,1}$	$\hat{\sigma}_{\hat{\theta},1}$	$\hat{\sigma}_{\psi,1}$	$\rho_{r\hat{\theta},1}$
0.60	0.124	−0.58	1.23	0.75	0.73	0.63	−0.34
0.70	0.124	−0.56	1.26	0.71	0.71	0.60	−0.34
0.80	0.125	−0.53	1.29	0.67	0.69	0.57	−0.34
0.90	0.125	−0.51	1.31	0.64	0.67	0.55	−0.34
1.00	0.126	−0.49	1.33	0.61	0.66	0.53	−0.34
1.10	0.126	−0.47	1.35	0.58	0.64	0.52	−0.34
1.20	0.127	−0.45	1.37	0.56	0.63	0.51	−0.35

Table 8: Angular Orientations for C-5R

Geometry	Ring	$\psi_{\ell,1}$	$\psi_{\ell,2}$	$\psi_{\ell,3}$	$\psi_{\ell,4}$	$\psi_{\ell,5}$
C-5R	$\ell = 1$	0.00π	0.40π	0.80π	1.20π	1.60π

individual ring components and that its weight decreases slightly as bias significance increases. It is seen that the covariance of the individual Gaussian components is decreased as the bias significance increases. Additionally, the components become less displaced in range ($\Delta\hat{r}_c$ and $\Delta\hat{r}_1$) and the ring components become more displaced in angle ($\Delta\hat{\theta}_1$). This shows that, for higher bias significance cases, the optimal solution tends to capture the overall distribution covariance using an increased displacement of the component means rather than increased component covariance.

A larger set of lookup tables for 3D GM model parameters are given in Appendix B. Not all geometries in Table 5 are listed in the appendix, but only those that were necessary to run the MAGMF for the scenarios described later in this chapter. Note that due to the measurement pre-conditioning technique used in the MAGMF, the need for GM lookup tables for a mixture of large size is greatly reduced. This is because after conditioning, the bias significance resulting from the state component cross-range variance (inflated by a small factor) and the measurement range variance drives the size of the mixture. The state components have a much more limited cross-range variance than the measurement. Therefore, only the initial measurement estimate (for filter initialization) requires the use of a large GM to achieve a low KL divergence fit.

4.4.2 Generating a Mixture from Normalized Lookup Parameters

Given a measurement with spherical means $\bar{r}$, $\bar{\alpha}$, and $\bar{\varepsilon}$ and standard deviations σ_r, σ_α, and σ_ε, a mixture model can be generated by the following process. First, compute the total bias significance C_B and bias significance ration R_B using (224) and (226).

(This section assumes that $R_B = 1$ for the measurement, a procedure to address the case where $R_B \neq 1$ is given in Section 4.5.2). For a given number of components N_G consult the lookup table for the mixture parameters evaluated at the measurement C_B.

If a central component is present in the mixture, the parameters ω_c, $\Delta\hat{r}_c$, $\hat{\sigma}_{r,c}$, $\hat{\sigma}_{\alpha,c}$, and $\hat{\sigma}_{\varepsilon,c}$ specify the weight, mean, and covariance of this component. The table parameters are first be converted a spherical normalized mean and covariance, according to

$$\Delta\hat{\mu}_{S,c} = \begin{bmatrix} \Delta\hat{r}_c \\ 0 \\ 0 \end{bmatrix} \tag{242}$$

and

$$\hat{\Sigma}_{S,c} = \begin{bmatrix} \hat{\sigma}^2_{r,c} & 0 & 0 \\ 0 & \hat{\sigma}^2_{\alpha\varepsilon,c} & 0 \\ 0 & 0 & \hat{\sigma}^2_{\alpha\varepsilon,c} \end{bmatrix} \tag{243}$$

All other components in the mixture belong to one or more rings. Let ℓ be the index of a given ring. Then the table parameters ω_ℓ, $\Delta\hat{r}_\ell$, $\Delta\hat{\theta}_\ell$, $\hat{\sigma}_{r,\ell}$, $\hat{\sigma}_{\theta,\ell}$, $\hat{\sigma}_{\psi,\ell}$, and $\rho_{r\theta,\ell}$, $\{\psi_{k_\ell}\}_{k_\ell \in I_\ell}$ specify the weight, mean, and covariance of the components in the ring, where $I_\ell \subseteq \{1, \ldots, N_G\}$ is the set of component indices associated with ring ℓ. Note that the only table parameter which is distinct for every component in the ring is ψ_{k_ℓ}. These table parameters are also converted to a normalized spherical mean and covariance, according to

$$\Delta\hat{\mu}_{S,k_\ell} = \begin{bmatrix} \Delta\hat{r}_\ell \\ \Delta\hat{\theta}_\ell \sin\psi_{k_\ell} \\ \Delta\hat{\theta}_\ell \cos\psi_{k_\ell} \end{bmatrix} \tag{244}$$

$$\hat{\Sigma}_{\theta\psi,k_\ell} = \begin{bmatrix} \hat{\sigma}^2_{r,\ell} & \hat{\sigma}_{r,\ell}\hat{\sigma}_{\theta,\ell}\rho_{r\theta,\ell} & 0 \\ \hat{\sigma}_{r,\ell}\hat{\sigma}_{\theta,\ell}\rho_{r\theta,\ell} & \hat{\sigma}^2_{\theta,\ell} & 0 \\ 0 & 0 & \hat{\sigma}^2_{\psi,\ell} \end{bmatrix} \tag{245}$$

$$\hat{\Sigma}_{S,k_\ell} = \begin{bmatrix} 1 & 0_{1\times 2} \\ 0_{2\times 1} & J_{\theta\psi,k_\ell}^{-T} \end{bmatrix} \hat{\Sigma}_{\theta\psi,k_\ell} \begin{bmatrix} 1 & 0_{1\times 2} \\ 0_{2\times 1} & J_{\theta\psi,k_\ell}^{-1} \end{bmatrix} \tag{246}$$

where

$$J_{\theta\psi,k_\ell}^{-1} = \begin{bmatrix} \sin\psi_{k_\ell} & \cos\psi_{k_\ell} \\ \hat{\sigma}_{\theta,\ell}\cos\psi_{k_\ell} & -\hat{\sigma}_{\theta,\ell}\sin\psi_{k_\ell} \end{bmatrix} \tag{247}$$

At this point all component means and covariances are represented in normalized spherical space ($\Delta\hat{\mu}_{S,k}$, $\hat{\Sigma}_{S,k}$), regardless of whether they originated as a central component or a ring component. Given the scaling matrices

$$S_\mu = \begin{bmatrix} \sigma_r\sqrt{2}C_B & 0 & 0 \\ 0 & \sigma_\alpha & 0 \\ 0 & 0 & \sigma_\varepsilon \end{bmatrix} \tag{248}$$

$$S_\Sigma = \begin{bmatrix} \sigma_r\sqrt{1+2C_B^2} & 0 & 0 \\ 0 & \sigma_\alpha & 0 \\ 0 & 0 & \sigma_\varepsilon \end{bmatrix} \tag{249}$$

convert to un-normalized space with

$$\begin{aligned} \mu_{S,k} &= \mu_S + S_\mu \Delta\hat{\mu}_{S,k} \\ \Sigma_{S,k} &= S_\Sigma \hat{\Sigma}_{S,k} S_\Sigma \end{aligned} \tag{250}$$

Finally, use (190) and (191) to compute the final Cartesian means and covariances.

$$\begin{aligned} \mu_k &= \mathrm{S2C}(\mu_{S,k}) \\ \Sigma_k &= J_{\mathrm{S2C}}(\mu_{S,k})\Sigma_{S,k}J_{\mathrm{S2C}}(\mu_{S,k})^T \end{aligned} \tag{251}$$

4.4.3 Adaptive Choice of Mixture Size Based on KL Divergence

In a similar way to Chapter 2, let a mixture of size N_G generated using this table lookup procedure for a spherical measurement with mean μ_S and covariance Σ_S be denoted by $\{\omega_k, \mu_k, \Sigma_k\}_{k=1}^{N_G} = \mathrm{Mix}_{\mathrm{EM}}(\mu_S, \Sigma_S, N_G)$, and the KL divergence of the mixture predicted by the lookup table be $\mathrm{KL}_{\mathrm{EM}}(\mu_S, \Sigma_S, N_G)$. If a target KL divergence

of $KL^\star$ is specified then let the minimum number of components needed to satisfy this limit be given by

$$N_G^\star(\mu_S, \Sigma_S, KL^\star) = \underset{N_G}{\arg\min}\, \mathrm{KL}_{\mathrm{EM}}(\mu_S, \Sigma_S, N_G) \leq KL^\star \tag{252}$$

Additionally, let

$$\mathrm{Mix}_{\mathrm{EM}}^\star(\mu_S, \Sigma_S, KL^\star) = \mathrm{Mix}_{\mathrm{EM}}(\mu_S, \Sigma_S, N_G^\star(\mu_S, \Sigma_S, KL^\star)) \tag{253}$$

and

$$\mathrm{KL}_{\mathrm{EM}}^\star(\mu_S, \Sigma_S, KL^\star) = \mathrm{KL}_{\mathrm{EM}}(\mu_S, \Sigma_S, N_G^\star(\mu_S, \Sigma_S, KL^\star)) \tag{254}$$

be the mixture and evaluated KL divergence which satisfies this limit. This procedure allows a filter to adaptively choose the number of components in a GM measurement model so that the maximum KL divergence of the model from the true distribution is limited.

4.5 Techniques to Generate GM Models for Bias Significance Ratios Other Than One

In this initial study, a lookup table was created to allow rapid modeling of measurements with unity bias significance ratios ($R_B = 1$) by a Gaussian mixture. However, this restriction on bias significance ratio results in major difficulties when applied to any practical sensor. Even sensors with equal azimuth and elevation errors $\sigma_\alpha = \sigma_\varepsilon$ have a unity bias significance ratio only when tracking at zero elevation angle. Therefore it is advantageous to provide a suboptimal solution for fits to measurements with bias significance ratios slightly different from $R_B = 1$. This section presents a process that can be used address this shortcoming. First, some preliminary techniques and concepts are described that will assist in describing the proposed approaches to $R_B \neq 1$ measurement modeling. Subsequently, two methods are presented that can be used to adapt $R_B = 1$ solutions to a alternative bias significance ratio. These KL divergence of GMs generated using these methods are then evaluated in comparison to the optimal KL divergence that can be achieved using ML.

4.5.1 Measurement Model Definitions

In a similar way to Section 2.4.1 in Chapter 2, it will be useful to define certain state-to-measurement transformation functions that may be used in a Kalman filter update according to the functional notation of Section 2.4.2. Let the Cartesian state to spherical measurement transformation be given by

$$h_S(x) = \text{C2S}([x]_p) \qquad H_S(x) = \begin{bmatrix} J_{\text{C2S}}([x]_p) & 0_{3\times 3} \end{bmatrix} \tag{255}$$

where $[x]_p$ indicates the position subvector of x, and the inverse of this function (with respect to the Cartesian position only) is expressed by

$$h_S^{-1}(z) = \text{S2C}(z) \tag{256}$$

For factorization of the angle marginal distributions, define the conversion functions of state onto angle as

$$\alpha = \text{C2A}\left(x_C\right) = \tan^{-1}\frac{x}{y} \tag{257}$$

$$J_{\text{C2A}}(x_C) = \begin{bmatrix} -\frac{y}{r_{xy}} & \frac{x}{r_{xy}} & 0 \end{bmatrix} \tag{258}$$

$$\varepsilon = \text{C2E}\left(x_C\right) = \sin^{-1}\frac{z}{\sqrt{x^2+y^2+z^2}} \tag{259}$$

$$J_{\text{C2E}}(x_C) = \frac{1}{x^2+y^2+z^2}\begin{bmatrix} -\frac{xz}{r_{xy}} & -\frac{yz}{r_{xy}} & r_{xy} \end{bmatrix} \tag{260}$$

where $r_{xy} = \sqrt{x^2+y^2}$. These angle transforms may be implemented in the Kalman filter as

$$h_\alpha(x) = \text{C2A}([x]_p) \tag{261}$$

$$H_\alpha(x) = \begin{bmatrix} J_{\text{C2A}}([x]_p) & 0 & 0 & 0 \end{bmatrix} \tag{262}$$

$$h_\varepsilon(x) = \text{C2E}([x]_p) \tag{263}$$

$$H_\varepsilon(x) = \begin{bmatrix} J_{\text{C2E}}([x]_p) & 0 & 0 & 0 \end{bmatrix} \tag{264}$$

If the measurements are converted to Cartesian space, then observation is simply a linear projection onto the position subspace.

$$h_C(x) = [x]_p = H_C x \tag{265}$$

$$H_C = \begin{bmatrix} I_{3\times 3} & 0_{3\times 3} \end{bmatrix} \tag{266}$$

4.5.2 Adaptation of Solutions for Bias Significance Ratios of One to Alternative Ratios

Two approaches were examined to allow the lookup tables for $R_B = 1$ measurements to be used to generate GMs for measurements with $R_B \neq 1$. The first approach uses the $R_B = 1$ normalized parameter lookup table and the transformations in Section 4.4.2 to generate a mixture using the true σ_α and σ_ε, despite the fact that the measurement does not have a bias significance ratio of exactly $R_B = 1$. The only difference is that the mean position scaling matrix from (248) is modified to use the general range shift scaling factor from (227) for $R_B \neq 1$. Let this approach be called the "rescaling" strategy.

$$S_\mu = \begin{bmatrix} \sigma_r \frac{1+R_B}{\sqrt{1+R_B^2}} C_B & 0 & 0 \\ 0 & \sigma_\alpha & 0 \\ 0 & 0 & \sigma_\varepsilon \end{bmatrix} \tag{267}$$

The second approach forms a measurement with the same C_B as the original measurement, but with $R_B = 1$. This is done by modifying the angle standard deviations only. Let $\sigma_{\alpha,0}$, and $\sigma_{\varepsilon,0}$ represent the standard deviations of this modified measurement. Using (226) and the constraint that C_B remains constant, solve for these values as follows

$$\sigma_{\varepsilon,0} = \sigma_\varepsilon \left(\frac{1+R_B^2}{2} \right)^{1/4} \quad \sigma_{\alpha,0} = \frac{\sigma_{\varepsilon,0}}{\cos \varepsilon} \tag{268}$$

Let $\mathrm{Diag}(x)$ denote a matrix formed by placing the elements of the vector x on it's diagonal with all other elements zero. Given a number of components N_G, generate

a baseline mixture using the lookup table method

$$\{\omega_{k,0}, \mu_{k,0}, \Sigma_{k,0}\}_{k=1}^{N_G} = \text{Mix}_{\text{EM}}(\mu_S, \text{Diag}\left(\begin{bmatrix}\sigma_r^2 & \sigma_{\alpha,0}^2 & \sigma_{\varepsilon,0}^2\end{bmatrix}\right), N_G) \tag{269}$$

Next, compute the difference in information in Az and El for this adjusted measurement.

$$\Delta I_\alpha = \frac{1}{\sigma_\alpha^2} - \frac{1}{\sigma_{\alpha,0}^2} \quad \Delta I_\varepsilon = \frac{1}{\sigma_\varepsilon^2} - \frac{1}{\sigma_{\varepsilon,0}^2} \tag{270}$$

If $R_B > 1$, then $\Delta I_\varepsilon > 0$ and $\Delta I_\alpha < 0$, which indicates that the modified measurement provides too little information in elevation and too much information in azimuth. If $R_B < 1$, the reverse is true. Due to this mismatch in information, the mixture result must be adjusted.

In the $R_B > 1$ case of too little information in elevation, a 1-D Kalman update may be performed in elevation with the mean $\bar{\varepsilon}$ and variance $1/\Delta I_\varepsilon$ (using the notation described in Section 2.4.2).

$$\begin{aligned}\mu'_{k,0} &= U_{x,\varepsilon}(\mu_{k,0}, \Sigma_{k,0}, \bar{\varepsilon}, \frac{1}{\Delta I_\varepsilon}) \\ \Sigma'_{k,0} &= U_{P,\varepsilon}(\mu_{k,0}, \Sigma_{k,0}, \bar{\varepsilon}, \frac{1}{\Delta I_\varepsilon})\end{aligned} \tag{271}$$

To update the weights, renormalize by the likelihood evaluated at the original mean and covariance according to

$$\omega'_{k,0} \propto \omega_{k,0} L_{x,\varepsilon}(\mu_{k,0}, \Sigma_{k,0}, \bar{\varepsilon}, \frac{1}{\Delta I_\varepsilon}) \tag{272}$$

In order to counteract the extra information in azimuth, an 1-D inverse Kalman update in azimuth may be performed based on the presence of an excess measurement with mean $\bar{\alpha}$ and variance $-1/\Delta I_\alpha$ (note $\Delta I_\alpha < 0$ in this case) according to

$$\begin{aligned}\mu_k &= U_{x,\alpha}^{-1}(\mu'_{k,0}, \Sigma'_{k,0}, \bar{\alpha}, -\frac{1}{\Delta I_\alpha}) \\ \Sigma_{k,0} &= U_{P,\alpha}^{-1}(\mu'_{k,0}, \Sigma'_{k,0}, \bar{\alpha}, -\frac{1}{\Delta I_\alpha})\end{aligned} \tag{273}$$

The inverse Kalman update process and associated notation are discussed in Section 2.4.3. To update the weights, renormalize by the inverse likelihood evaluated at

the updated mean and covariance.

$$\omega_k \propto \frac{\omega'_{k,0}}{L_{x,\varepsilon}(\mu_k, \Sigma_k, \bar{\alpha}, -\frac{1}{\Delta I_\alpha})} \tag{274}$$

For the $R_B < 1$ case, simply exchange α and ε in (271)-(274) and perform Kalman update in azimuth and inverse update in elevation. Let this approach be called the "conditioning" strategy.

In order to test the effectiveness of these approaches, $N_G = 6$ C-5R fits to a measurement with $\bar{\varepsilon} = 0$ rad, $\sigma_\varepsilon = 1$ mrad, $\bar{r} = 100$ km, and $C_B = 1.5$ were attempted using both the EM algorithm and the mixture adaptation procedures described above. The ratio of azimuth to elevation error (equivalent to $\sqrt{R_B}$) was varied logarithmically from 0.5 to 2.0 and used to set the azimuth error σ_α.

Figure 101 shows the results of this study. The yellow line "Lookup for Same C_B, $R_B = 1$" indicates the baseline KL divergence predicted by the lookup table for a C-5R fit to $C_B = 1.5$ and $R_B = 1$. The goal is that the adapted solutions not perform significantly worse than this in terms of KL divergence. The purple line "EM Fit" indicates the performance of the EM algorithm in fitting a mixture of size $N_G = 6$ to the measurement in question based on an initialization on the C-5R solution. When $R_B \neq 1$ a lower KL divergence is possible as seen in Section 4.2, as EM is able to modify the geometry of the components in the solution to more appropriately fit the measurement in question.

The red line "Conditioned Mixture" indicates the KL divergence achieved by using the mixture conditioning approach described in this section. From about $\sqrt{R_B} = 0.8$ to $\sqrt{R_B} = 1.2$ these mixtures have KL divergence performance very close to that of a measurement with the same C_B but $R_B = 1$. Outside of this interval, the KL divergence penalty is approximately 5%.

Finally, the blue line "Rescaled Mixture" shows the results based on the rescaling approach where the $R_B = 1$ lookup tables are simply used with appropriate scaling by the true Az and El error standard deviations. Over the region of study, it is seen

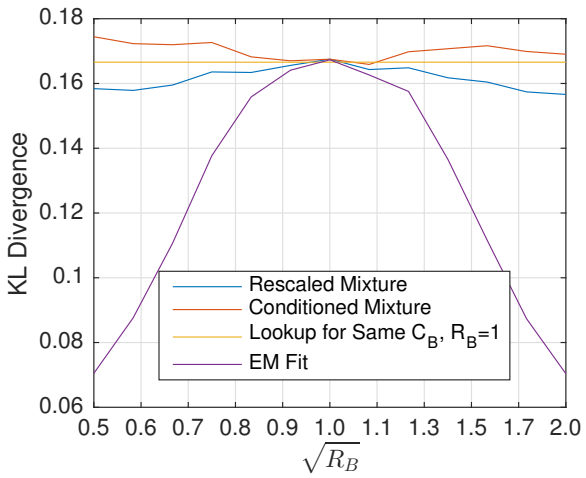

Figure 101: KL Divergence for Adapted C-5R Solutions for $C_B = 1.5$

that the mixtures generated by this approach have equal or lesser KL divergence than the baseline $R_B = 1$ case, with the maximum improvement begin approximately 5%. Therefore, this approach is preferred to the conditioning approach due to its superior performance and simplicity.

Overall, the rescaled adapted solutions perform significantly worse than the optimal KL divergence achievable by adopting the EM solution for a given R_B. However, they provide an acceptable alternative to creating additional lookup tables for solutions with $R_B \neq 1$ if one is willing to use the same number of components to model these cases as the $R_B = 1$ case, since the performance will not be degraded from this baseline.

4.6 Estimation and Tracking Using GM Measurement Models

In order to evaluate the utility of this 3D measurement modeling procedure for range-az-el measurements, it will be used as part of a Gaussian mixture Kalman filter [1]. The architecture of this filter is based on that of the 2D MAGMF described in Chapter 2. The performance of this 3D MAGMF is compared to existing filtering techniques that have been proposed for addressing the 3D contact-lens problem. These

include

1. Standard EKF with Cartesian state and polar measurement
2. UKF with Cartesian state and polar measurement [16]
3. MCAEKF [27]
4. GMM-ITS [32]
5. Regularized Particle Filter (RPF) [24]

The implementation of these alternative filters is described in detail in Chapter 2 and their respective references. For the MCAEKF, the bias significance in (224) is limited to $C_B \leq 0.2$ - above this level, range variance inflation occurs. The GMM-ITS was run with a $N_G = 4 \times 4$ split in Az and El and $\alpha = 3$ as described in [33] for the longer range scenario. The pruning threshold for low weight components was set to 10^{-6}. For the RPF, $300,000$ particles and $h = 2$ Epanechnikov kernel bandwidth were used as suggested by [33] for the 3D scenarios tested in that work. An effective sample threshold of $150,000$ was used.

4.6.1 Filter Dynamics Model

For all scenarios run here, the dynamics model used is nearly-constant-velocity (NCV) with small process noise $Q_c = 10^{-3} I_3$ or $Q_c = 10^{-4} I_3$, depending the scenario (where I_3 is an identity matrix in 3D). This linear model has the following mean and covariance update equations which can be found in standard references [2].

$$\hat{x}_k = \begin{bmatrix} \hat{x}_{1,k} & \hat{x}_{2,k} & \hat{x}_{3,k} & \hat{\dot{x}}_{1,k} & \hat{\dot{x}}_{2,k} & \hat{\dot{x}}_{3,k} \end{bmatrix}^T \tag{275}$$

$$F(\Delta t) = \begin{bmatrix} I_3 & \Delta t I_3 \\ 0_3 & I_3 \end{bmatrix} \tag{276}$$

$$Q(\Delta t) = \begin{bmatrix} \frac{1}{3}Q_c\Delta t^3 & \frac{1}{2}Q_c\Delta t^2 \\ \frac{1}{2}Q_c\Delta t^2 & Q_c\Delta t \end{bmatrix} \tag{277}$$

$$\hat{x}_{k|k-1} = F(\Delta t)\hat{x}_{k-1|k-1} \tag{278}$$

$$P_{k|k-1} = F(\Delta t)P_{k-1|k-1}F(\Delta t)^T + Q(\Delta t) \tag{279}$$

Here Δt is the filter propagation time-step and $\hat{x}_{(\cdot),k}$ and $\hat{\dot{x}}_{(\cdot),k}$ are the scalar components of the state vector representing 3D Cartesian position and velocity, respectively.

4.6.2 Filter Initialization

All parametric filters were initialized using the so-called "one-point" initialization technique. For single-Gaussian filters, the position portion of $\hat{x}_{0|0}$ and $P_{0|0}$ are set to the mean and covariance of the converted measurement z_0 and R_0. For the Gaussian mixture filters (MAGMF, GMM-ITS), a mixture representation $\omega_{0|0,i}$, $\hat{x}_{0|0,i}$, and $P_{0|0,i}$ of the position distribution is generated by the respective techniques used in the filter. The velocity portion of the state is initialized to a mean of zero and a diagonal covariance of $\frac{v_{max}^2}{3}$, where $v_{max} = 5000$ m/s is a prior knowledge parameter on the maximum speed of the target.

For the RPF, the position components of the initial particles $x_{0|0,i}$ are drawn from the Gaussian measurement distribution in spherical space and these samples then converted to Cartesian space by way of the spherical-to-Cartesian transformation (190). The velocity samples are then generated by the two-point differencing method described in [24]. All initial weights $\omega_{0|0,i}$ are set uniformly to $\frac{1}{N_p}$, where N_p is the number of particles.

4.6.3 3D MAGMF Implementation

The 3D MAGMF operates very similarly to the 2D MAGMF in Chapter 2. However, some initial care must be taken during the measurement pre-conditioning step to

handle spherical measurements.

The filter is initialized with a mixture by the one-point method described in Section 4.6.2, where the initial position mixture is given by $\text{Mix}^{\star}_{\text{EM}}(z_0, R_0, KL^{\star})$.

As in the 2D case, for each time step k, the previous posterior

$$\left\{\omega_{x,k-1|k-1,i}, \hat{x}_{k-1|k-1,i}, P_{k-1|k-1,i}\right\}_i$$

is updated to the current prior

$$\left\{\omega_{x,k|k-1,i}, \hat{x}_{k|k-1,i}, P_{k|k-1,i}\right\}_i$$

by independent component-wise application of the dynamics in (278) and (279) with no modification to the weights.

The update procedure for each component i is as follows. As in the 2D case, evaluate the equivalent state azimuth and elevation means and covariances.

$$\begin{aligned} \alpha_{x_{k|k-1},i} &= h_{\alpha}(x_{k|k-1,i}) \quad \varepsilon_{x_{k|k-1},i} = h_{\varepsilon}(x_{k|k-1,i}) \\ \sigma^2_{\alpha,x_{k|k-1},i} &= H_{\alpha}(\hat{x}_{k|k-1,i}) P_{k|k-1,i} H_{\alpha}(\hat{x}_{k|k-1,i})^T \\ \sigma^2_{\varepsilon,x_{k|k-1},i} &= H_{\varepsilon}(\hat{x}_{k|k-1,i}) P_{k|k-1,i} H_{\varepsilon}(\hat{x}_{k|k-1,i})^T \end{aligned} \tag{280}$$

Compute the total bias significance (224) of the measurement compared to the state using

$$C_B = \frac{r_k \sqrt{\sigma^4_{\varepsilon,x_{k|k-1},i} + \sigma^4_{\alpha,x_{k|k-1},i} \cos^4(\varepsilon_{x_{k|k-1},i})}}{2\sigma_{r,k}} \tag{281}$$

where r_k and $\sigma_{r,k}$ are the range and range standard deviation of the measurement. Then, use (222) to compute the equivalent single-Gaussian KL-divergence of the model. If this is less than the chosen limit of $KL^{\star}$, update the track component in polar space using the MCAEKF technique (using the notation of Section 2.4.2 along with the measurement model defined in Section 4.5.1).

$$\hat{x}_{k|k,i} = U_{x,S}(\hat{x}_{k|k-1,i}, P_{k|k-1,i}, z_k, R'_{k,i}) \tag{282}$$

$$P_{k|k,i} = U_{P,S}(\hat{x}_{k|k-1,i}, P_{k|k-1,i}, z_k, R'_{k,i}) \tag{283}$$

where $R'_{k,i}$ is the spherical measurement covariance with range variance inflated using MCAEKF so that the bias significance limits the KL divergence to $KL^\star$.

If the single-Gaussian bias significance limit for MCAEKF is not met, then a conditioned Cartesian measurement mixture will be used to update the filter. With a given choice of κ, generate the conditioning measurement $\theta_{y,k,i}$ and $\sigma^2_{\theta,y,k,i}$ such that the angle mean and covariance of the measurement are updated to $\theta_{z'_k,i} = \theta_{x_{k|k-1},i}$ and $\sigma^2_{\theta,z'_k,i} = \kappa\sigma^2_{\theta,x_{k|k-1},i}$. Then use the lookup table to generate the mixture

$$\left\{\omega'_{z_C,k,i,j}, z'_{C,k,i,j}, R'_{C,k,j}\right\}_j = \mathrm{Mix}^\star_{\mathrm{EM}}(z'_{k,i}, R'_{k,i}, KL^\star)$$

where $z'_{k,i}$ and $R'_{k,i}$ are the measurement mean and covariance conditioned in angle to the support of state component i. The size of this mixture (chosen so that the KL divergence meets the limit $KL^\star$) will in general be smaller than generated by the unconditioned version of the process $\mathrm{Mix}^\star_{\mathrm{EM}}(z_k, R_k, KL^\star)$ due to the smaller angular variance. Update the prior component i with this mixture using

$$\hat{x}'_{k|k,i,j} = U_{x,C}(\hat{x}_{k|k-1,i}, P_{k|k-1,i}, z'_{C,k,i,j}, R'_{C,k,j}) \tag{284}$$

$$P'_{k|k,i,j} = U_{P,C}(\hat{x}_{k|k-1,i}, P_{k|k-1,i}, z'_{C,k,i,j}, R'_{C,k,j}) \tag{285}$$

and finally apply the inverse filtering

$$\hat{x}_{k|k,i,j} = U^{-1}_{x,\theta}(\hat{x}'_{k|k,i,j}, P'_{k|k,i,j}, \theta_{y,k,i}, \sigma^2_{\theta,y,k,i}) \tag{286}$$

$$P_{k|k,i,j} = U^{-1}_{P,\theta}(\hat{x}'_{k|k,i,j}, P'_{k|k,i,j}, \theta_{y,k,i}, \sigma^2_{\theta,y,k,i}) \tag{287}$$

to get the final posterior means and distributions of the state component i. Finally, find the weights of the posterior PDF conditioned on prior component i using

$$\tilde{\omega}_{x_{k|k},i,j} \propto \frac{\omega'_{z_C,k,i,j}}{L_{z,\theta}\left(\hat{x}_{k|k,i,j}, P_{k|k,i,j}, \theta_{y,k,i}, \sigma^2_{\theta,y,k,i}\right)} \tag{288}$$

normalized over j.

At this point, for every prior component i, a posterior conditional mixture $\left\{\tilde{\omega}_{x_{k|k},i,j}, \hat{x}_{k|k,i,j}, P_{k|k,i,j}\right\}_j$ has been calculated. (If MCAEKF was used to update prior

component i, then this mixture is trivially a single Gaussian with $\tilde{\omega}_{x_{k|k},i,1} = 1$). To find the final weights $\omega_{x_{k|k},i,j}$, update by the prior mixture weights and likelihoods of the measurement given the prior according to

$$\omega_{x_{k|k},i,j} \propto \tilde{\omega}_{x_{k|k},i,j}\omega_{x_{k|k-1},i}L_{x,P}\left(\hat{x}_{k|k-1,i}, P_{k|k-1,i}, z_k, R_k\right) \tag{289}$$

This is normalized to unity over all i, j. Note that the likelihood evaluation is performed in polar space using the original unmodified polar measurements. A Cartesian mixture should not be used for this purpose as it does not represent the likelihood well at the edges of the distribution.

As a final step, the multiplicity control algorithm described in Section 2.5.6.2 is applied. The only difference is that the equivalent spherical position distribution $\hat{z}_{k|k} = h_S(\hat{x}_{k|k})$ and $\hat{R}_{k|k} = H_S(\hat{x}_{k|k})P_{k|k}H_S(\hat{x}_{k|k})^T$ is used to calculate the target number of components using the 3D lookup table KL divergence according to

$$N^{\star}_{G,k|k} = N^{\star}_G(\hat{z}_{k|k}, \hat{R}_{k|k}, KL^{\star}_{pos}) \tag{290}$$

4.7 Simulation Results

The scenarios and radar parameters that will be tested are based on the scenarios in [33]. The radar measurement accuracy in range is $\sigma_r = 1$ m, and in angle $\sigma_\alpha = \sigma_\varepsilon = 1$ mrad. (The radar in the original paper operates in RUV space, but this adaptation to Range-Az-El is comparable). Three scenarios Z1, Z2, and Z3 with varying bias significance will be studied. The initial velocity for each of the scenarios is $\begin{bmatrix}-2 & -2 & -1\end{bmatrix}^T$ km/s, and the initial positions are $\begin{bmatrix}924 & 900 & 1050\end{bmatrix}^T$ km for Z1, $\begin{bmatrix}3080 & 3000 & 3500\end{bmatrix}^T$ km for Z2, and $\begin{bmatrix}3277 & 3192 & 0\end{bmatrix}^T$ km for Z3. Scenario Z1 has a total bias significance of $C_B = 0.97$ while Z2 and Z3 have a total bias significance of $C_B = 3.2$. Scenario Z1 and Z2 have a bias significance ratio of $R_B = 0.6$ while Z3 has $R_B = 1$. (In the original work [33], the short range scenario had a third of the range, which results in a bias significance of 0.3. This situation can be readily handled by a

single Gaussian model, so the range of this scenario is inflated for Z1 here. Scenario Z3 is not in the original work but represents an adaptation of Z2 to zero elevation angle so that an $R_B = 1$ case can be studied.) This mixture of scenarios allows study of the filter performance under a variety of different measurement geometries.

For each scenario, 100 Monte Carlo runs of each filter were performed in order to collect performance statistics. For the MAGMF, a KL divergence performance limit of $KL^{\star} = 0.05$ was set for the measurement and $KL^{\star}_{pos} = 0.0125$ for the number of posterior track components. The measurement conditioning covering multiplier was set to $\kappa = 1.5$.

For the MCAEKF, the bias significance in (224) is limited to $C_B \leq 0.2$ - above this level, range variance inflation occurs. The Zhang GMM-ITS was run with $\alpha = 3$ for the support coverage parameter and $N_G = 3 \times 3$ components for Z1, and $N_G = 4 \times 4$ components for Z2 and Z3, as in [33]. The pruning threshold for low weight components was set to 10^{-6}. For the RPF, 300k particles were used along with a Epanechnikov kernel bandwidth of $h = 0.7$ and an effective sample threshold of 150k particles. In [24] a bandwidth of $h = 2$ was given, but this results in an extremely pessimistic covariance and limited convergence in cross-range accuracy.

4.7.1 Scenario Z1 Filtering Results

Figure 102 shows the RMSE in range for Scenario Z1. As expected, the UKF, MCAEKF and standard EKF have degraded range performance during the initial 10 iterations of the filter. The MAGMF, GMM-ITS, and RPF are able track the range performance of the sensor at the start of the track during this initial period. However, the GMM-ITS has slightly worse performance due to the inflated range standard deviation used in it's mixture process ($\sigma_{r,i} = \alpha\sigma_r$). Figures 103-105 show that all filters besides the standard EKF have good RMSE performance in Az, El, and total position. Figure 107, shows that the same is true for total velocity RMSE.

However, Figure 106 shows that in range rate the MAGMF, GMM-ITS, and RPF have a superior convergence rate compared to the single-Gaussian filters.

For covariance consistency, the NEES is shown in Figure 108 based on a single-Gaussian moment match in spherical coordinate space for the position only and based on a single-Gaussian moment match in Cartesian coordinates for the combined position and velocity in Figure 109. In both cases, the NEES is not normalized by the state dimension, so the nominal values are 3 and 6 for the position and total state, respectively. The dashed lines show the 99% confidence region of the NEES based on the number of MC runs. In the mixture filters and particle filter, interpretation of the total NEES should be caveated as the true distribution in Cartesian space is non-Gaussian and the metric of NEES (Mahalanobis distance) is only strictly applicable to a Gaussian distribution. The distribution of the position state in spherical space is expected to more closely resemble a Gaussian. Thus NEES in spherical space may be more indicative of good performance.

Overall, the GMM-ITS has a pessimistic NEES in both spherical and Cartesian space due to the covariance inflation factor mentioned previously. The RPF is slightly pessimistic compared to the other filters due to the injected kernel bandwidth. The standard EKF is highly inconsistent as expected. For Scenario Z1, the UKF, MCAEKF and MAGMF all have excellent consistency.

Finally, a computational resource comparison between the GMM-ITS and MAGMF is presented for Z1. Figure 110 shows the average number of state components that must be stored for the track at a given time and Figure 111 shows the average number of total Kalman updates needed between state and measurement (the number of posterior components before component reduction is performed). Note that the UKF, EKF, and MCAEKF have an equivalent cost of 1 in both categories, while the RPF has a vastly higher cost than either of these filters in terms of both computational and storage resources. In terms of storage, the MAGMF starts with an average of

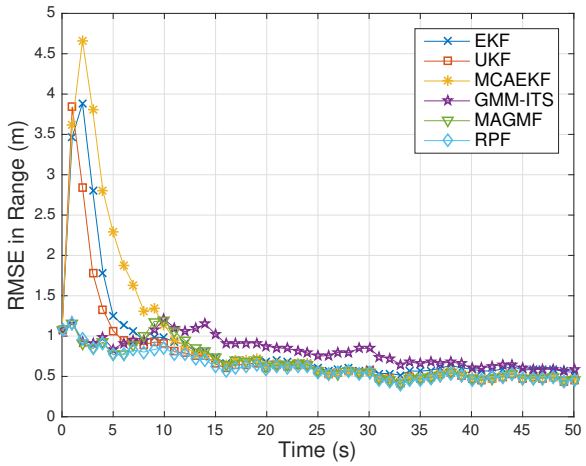

Figure 102: RMSE in Range for Scenario Z1 ($C_B = 1$, $R_B = 0.6$)

13 components (the maximum allowed by the lookup table used in this particular implementation of the filter) compared to the 9 used by GMM-ITS, but this quickly falls off adaptively as the size of the angular support region decreases, culminating in a single-Gaussian filter after approximately 17 s. In contrast, the number of computations needed to update the state mixture by the measurement mixture remains near 9^2 for almost the entire duration of the GMM-ITS due to the lack of adaptation, while it falls off rapidly in the MAGMF due to convergence. By $t = 10$ s the MAGMF is able to operate using only a handful of updates; essentially each state component is updated using the MCAEKF in this regime. However, if the filtering conditions change (such as due to missed measurements or increased process noise), the MAGMF can add more GMs to the model as needed.

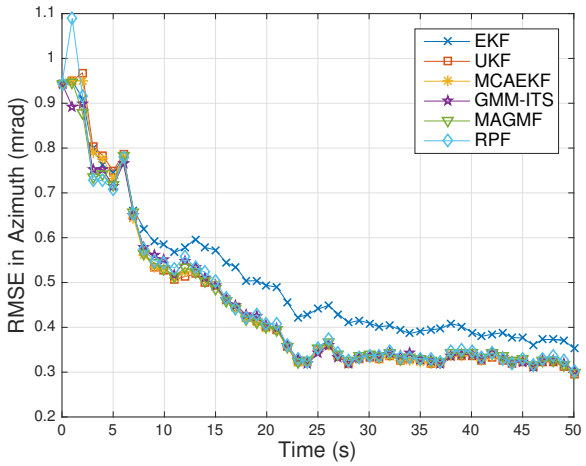

Figure 103: RMSE in Az for Scenario Z1 ($C_B = 1$, $R_B = 0.6$)

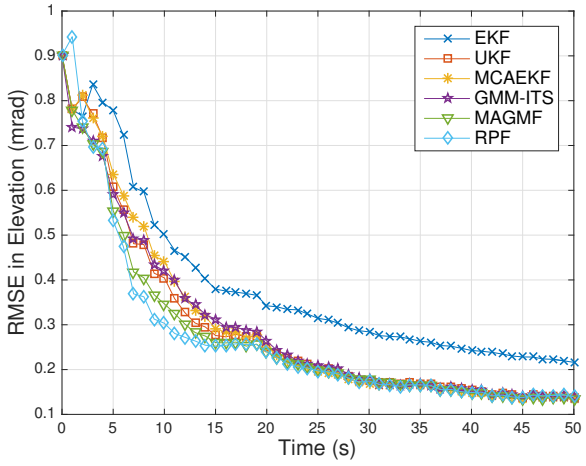

Figure 104: RMSE in El for Scenario Z1 ($C_B = 1$, $R_B = 0.6$)

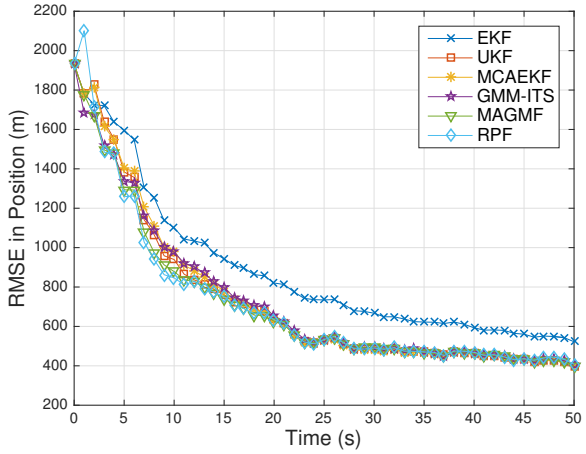

Figure 105: RMSE in Total Position for Scenario Z1 ($C_B = 1$, $R_B = 0.6$)

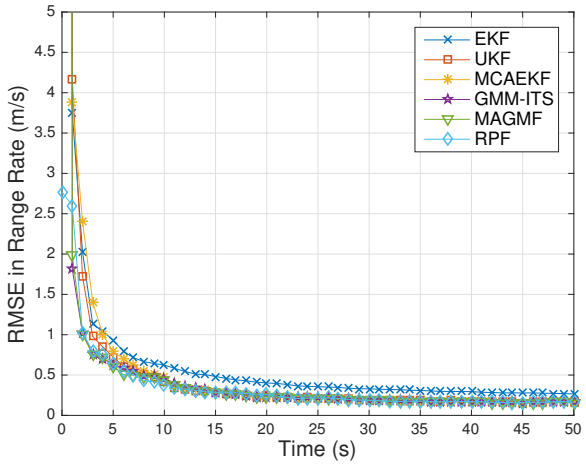

Figure 106: RMSE in Range Rate for Scenario Z1 ($C_B = 1$, $R_B = 0.6$)

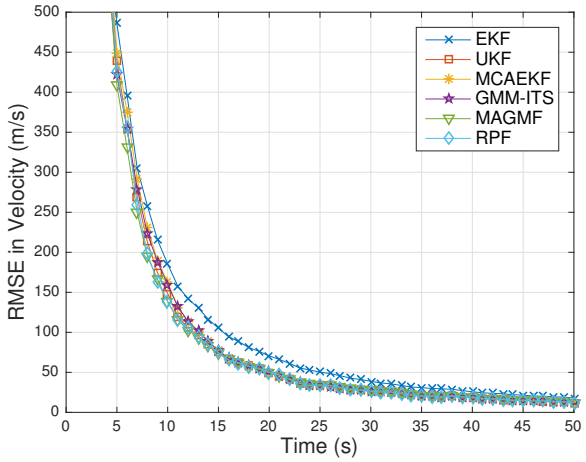

Figure 107: RMSE in Total Velocity for Scenario Z1 ($C_B = 1$, $R_B = 0.6$)

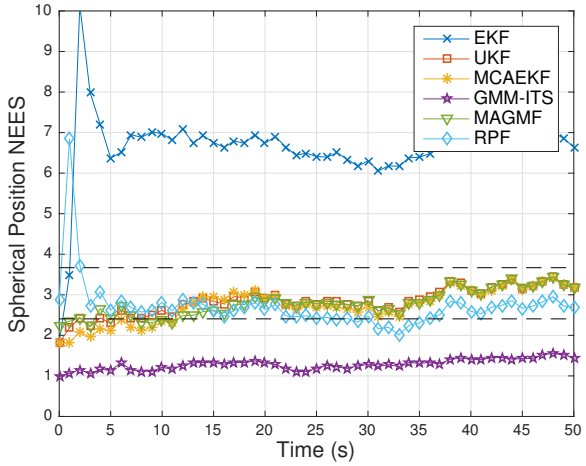

Figure 108: Spherical Position NEES for Scenario Z1 ($C_B = 1$, $R_B = 0.6$)

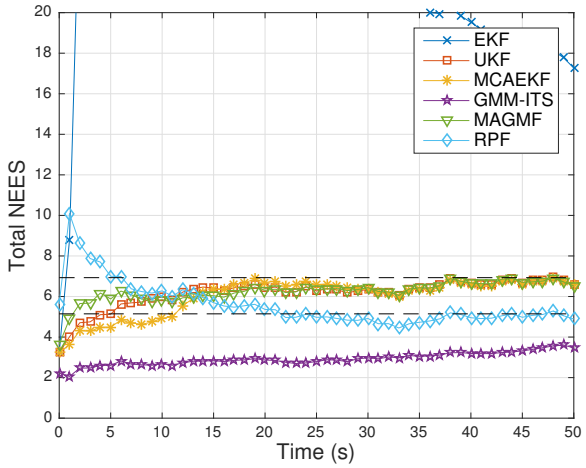

Figure 109: Total NEES for Scenario Z1 ($C_B = 1$, $R_B = 0.6$)

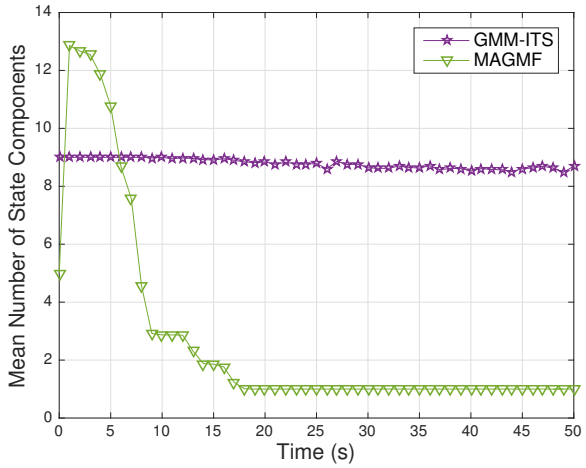

Figure 110: Number of State Components for Scenario Z1 ($C_B = 1$, $R_B = 0.6$)

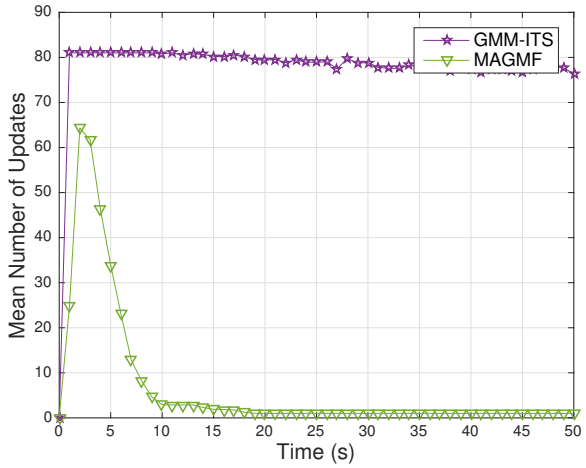

Figure 111: Number of Kalman Updates for Scenario Z1 ($C_B = 1$, $R_B = 0.6$)

4.7.2 Scenario Z2 Filtering Results

For Scenario Z2, the bias significance is greatly increased compared to Z1. Therefore, the EKF, UKF and MCAEKF exhibit degraded range estimation performance compared to the Z1 case, as shown in Figure 112. Additionally, in this case the GMM-ITS is not able to converge in range accuracy beyond the precision of the measurements due to the inflated range variance in the measurement mixture. The MAGMF and RPF are able to both perform at the precision of the measurements from the start of the track and converge to the same level of performance of the MCAEKF and UKF in the long term.

In terms of angle errors and total error performance, all filters except the standard EKF have similar performance for Z2, as shown in Figures 113 through 115. However, the MAGMF slightly outperforms the other filters at the beginning of the track, especially around $t = 5 - 10$ s in Az and $t = 15 - 20$ s in El. Similar results exist for the total velocity error in Figure 117 for Z2 where the MAGMF and RPF slightly outperform the other filters from $t = 10$ to 20 s. In range rate, shown in Figure 116, the GMM-ITS, RPF, and MAGMF converge significantly more quickly than the single-Gaussian filters.

In terms of covariance consistency, the RPF, MAGMF and MCAEKF have the best spherical position NEES shown in Figure 118, followed by the UKF. The EKF is extremely inconsistent (off the axes), while the GMM-ITS is pessimistic. If the total state NEES is examined in Figures 119 it appears the MAGMF and MCAEKF are inconsistent. However as previously stated, the PDF of the state in Cartesian space for $C_B = 3$ is highly non-Gaussian, so care must be taken when applying the NEES metric.

Finally, the computational costs of the MAGMF and GMM-ITS are shown in Figures 120 and 121. The number of MAGMF state components in Figure 120) remains close to the maximum of 13 near the start of the track, and begins to rapidly

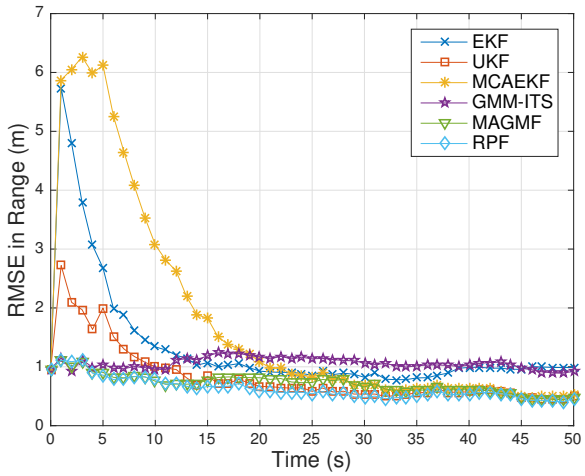

Figure 112: Range RMSE for Scenario Z2($C_B = 3$, $R_B = 0.6$)

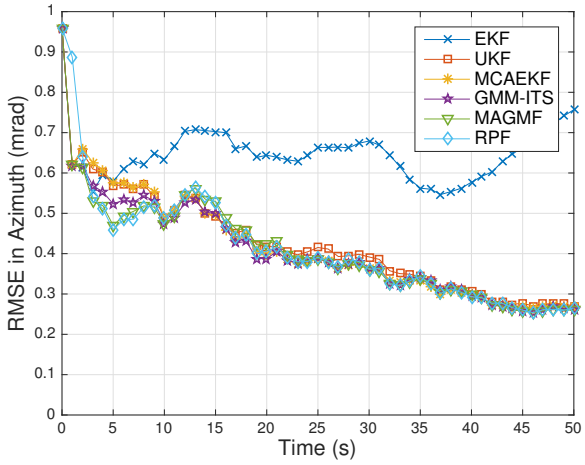

Figure 113: Az RMSE for Scenario Z2 ($C_B = 3$, $R_B = 0.6$)

converge around $t = 18$ s. The final average is $N_G = 2$ state components. An even more significance computational savings is evident in terms of the number of measurement-to-state Kalman updates required for the MAGMF vs the GMM-ITS. Figure 121 shows that the maximum average number of updates for the MAGMF is around 160 and rapidly falls off in the first few seconds, compared to the almost 250 updates for the GMM-ITS, which has a much slower rate of reduction due to pruning of low weight solutions.

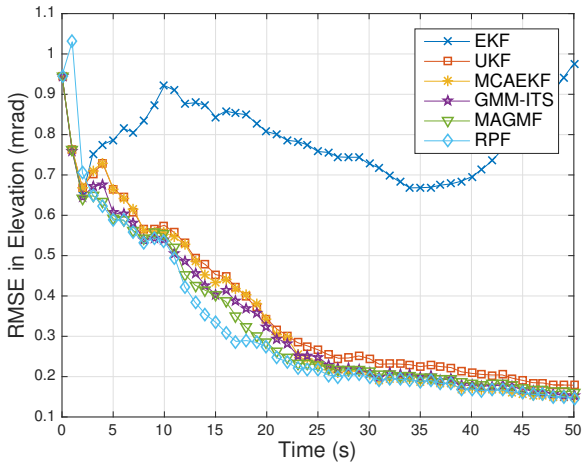

Figure 114: El RMSE for Scenario Z2 ($C_B = 3$, $R_B = 0.6$)

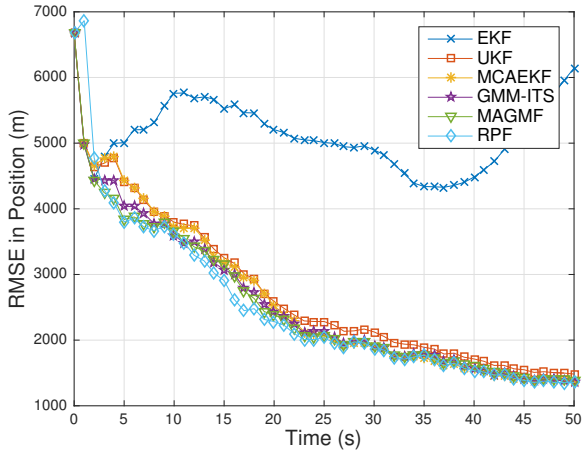

Figure 115: Total Position RMSE for Scenario Z2 ($C_B = 3$, $R_B = 0.6$)

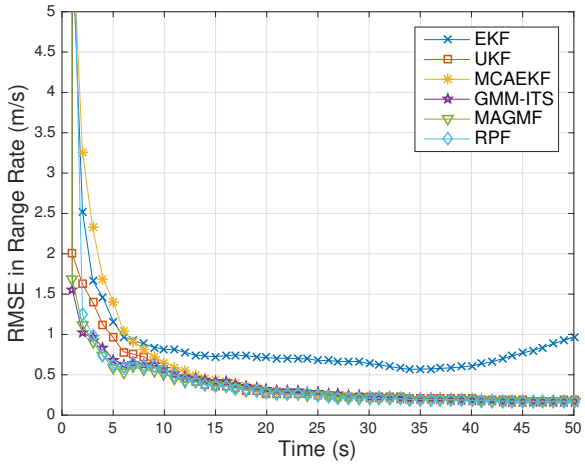

Figure 116: Range Rate RMSE for Scenario Z2 ($C_B = 3$, $R_B = 0.6$)

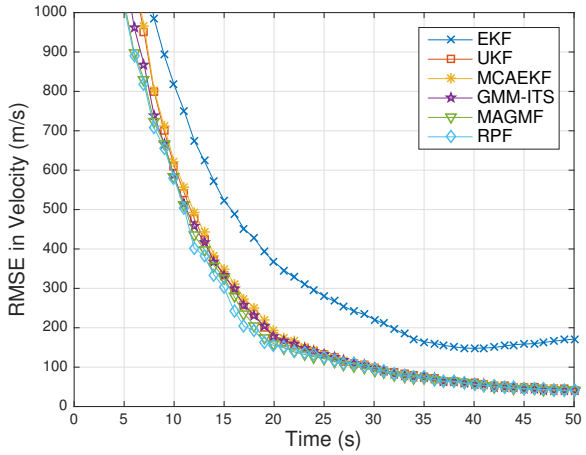

Figure 117: Total Velocity RMSE for Scenario Z2 ($C_B = 3$, $R_B = 0.6$)

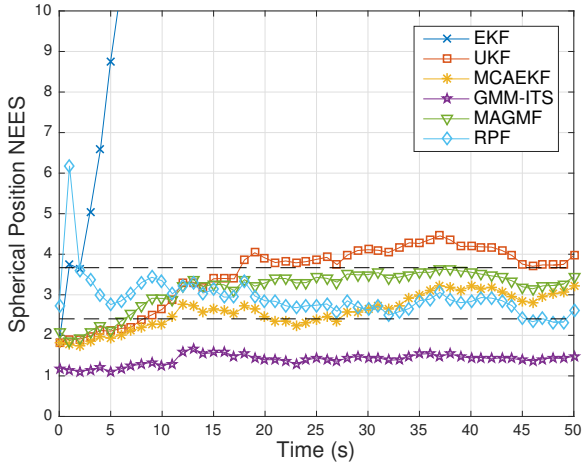

Figure 118: Spherical Position NEES for Scenario Z2 ($C_B = 3$, $R_B = 0.6$)

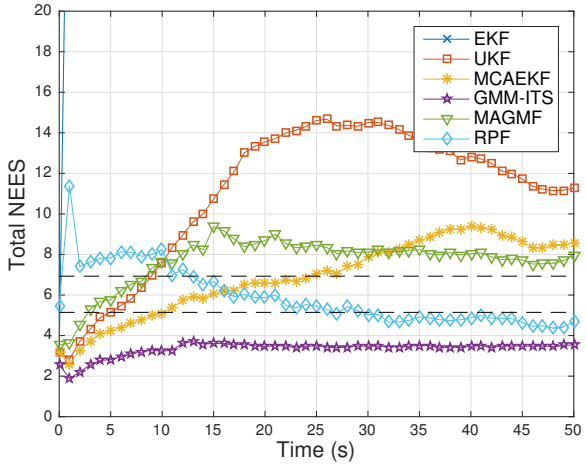

Figure 119: Total NEES for Scenario Z2 ($C_B = 3$, $R_B = 0.6$)

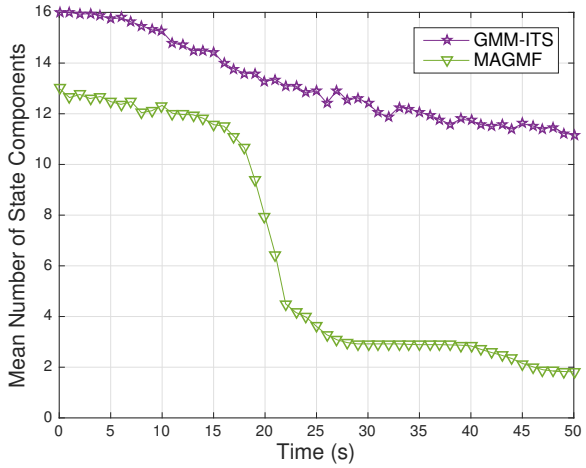

Figure 120: Number of State Components for Scenario Z2 ($C_B = 3$, $R_B = 0.6$)

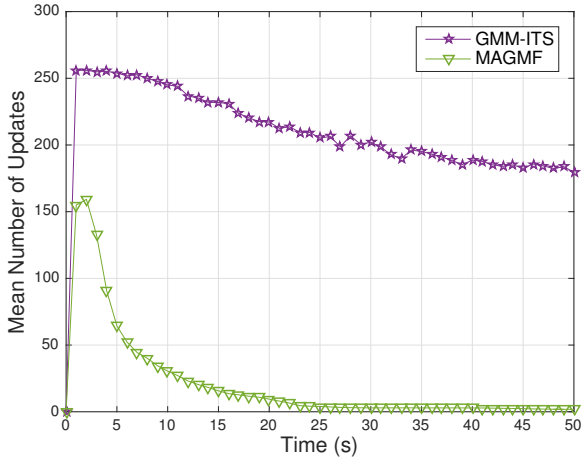

Figure 121: Number of Kalman Updates for Scenario Z2 ($C_B = 3$, $R_B = 0.6$)

4.7.3 Scenario Z3 Filtering Results

The results for Scenario Z3 are very similar to those of Z2, as the only difference is the change in bias significance ratio to $R_B = 1$. Figure 122 shows that the EKF, UKF and MCAEKF are not able to track with RMSE comparable to the measurements early in the track. The MAGMF and RPF have the best performance, and the GMM-ITS has slightly degraded range estimates in comparison. Figure 123 shows that in azimuth, the MCAEKF, RPF, and GMM-ITS have the best performance, while the EKF performs much worse than the other filters. The situation is similar in elevation as shown in Figure 124, except that the RPF and MAGMF have a clear edge in tracking performance compared to the GMM-ITS and single-Gaussian filters around $t = 15$ s. Figures 125 through 127 show a similar trend, with the RPF and MCAEKF having the best performance, followed by the GMM-ITS, the MCAEKF and UKF, and finally the EKF, which has very poor performance.

For covariance consistency, Figure 128 shows that the spherical NEES of the EKF is highly inconsistent and the GMM-ITS is slightly pessimistic. All other filters have reasonable spherical NEES. In total NEES, Figure 129 shows that the MAGMF and RPF have the closest to optimal NEES, with other filters inconsistent to some degree. However, as previously mentioned, care must be taken in interpreting the total NEES metric as it is based on a Gaussian moment match to the state, where the true state distribution is non-Gaussian.

Finally, the computational costs of the MAGMF and GMM-ITS are shown in Figures 130 and 131 Again, $12 - 13$ state components are used in Figure 130 until the filter begins to converge around $t = 22$ s. The final average is $N_G = 3$ state components. The difference in the convergence times between Z2 and Z3 is likely due to the more challenging geometry associated with the $R_B = 1$ measurements in Z3 compared to $R_B = 0.6$ in Z2. Again, the GMM-ITS begins with $N_G = 16$ components and converges slowly to a final size of $N_G = 12$ state components.

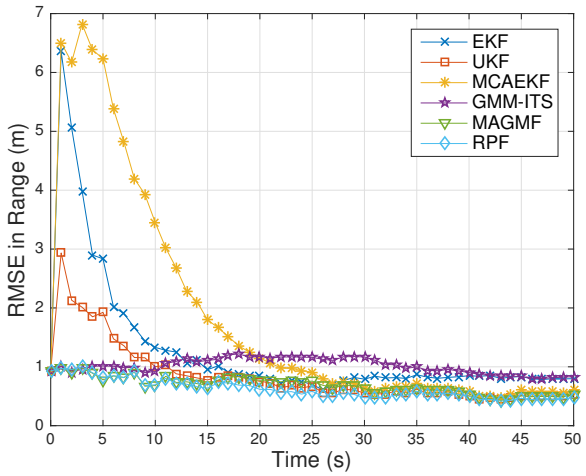

Figure 122: Range RMSE for Scenario Z3 ($C_B = 3$, $R_B = 1$)

Figure 131 shows the number of Kalman updates for the GMM-ITS and MAGMF for Z3. This result is very similar to the Z2 results in Figure 121. Again, the GMM-ITS requires a larger number of state-to-measurement updates in order to provide effective estimation performance compared to the MAGMF, which converges rapidly in number of updates over time.

Overall, the performance for Scenario Z3 is very similar to Z2 despite the difference in R_B. This is expected due to the use of the lookup table adaptation method described in Section 4.5.2 which generate mixtures for $R_B \neq 1$ that have a similar KL divergence to the $R_B = 1$ case. If additional lookup tables were created for $R_B \neq 1$, it is likely that fewer components and computations would be needed for Scenario Z2.

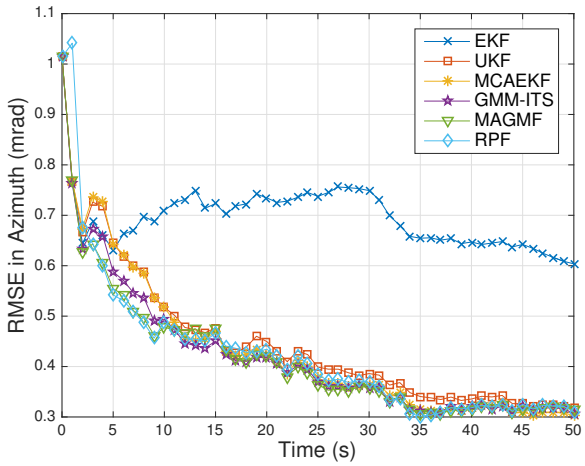

Figure 123: Az RMSE for Scenario Z3 ($C_B = 3$, $R_B = 1$)

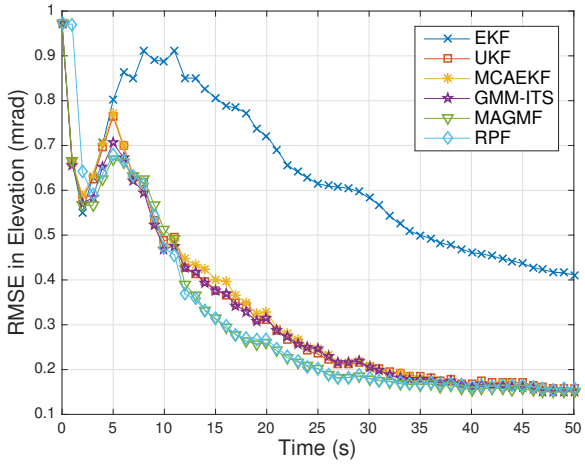

Figure 124: El RMSE for Scenario Z3 ($C_B = 3$, $R_B = 1$)

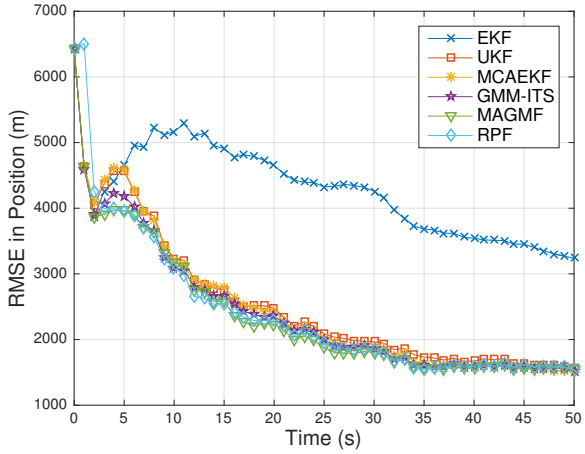

Figure 125: Total Position RMSE for Scenario Z3 ($C_B = 3$, $R_B = 1$)

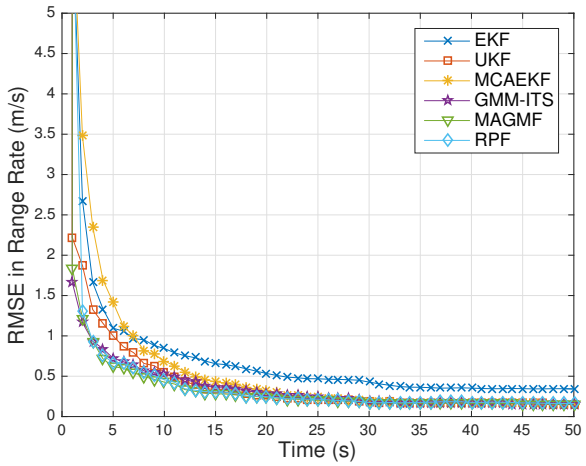

Figure 126: Range Rate RMSE for Scenario Z3 ($C_B = 3$, $R_B = 1$)

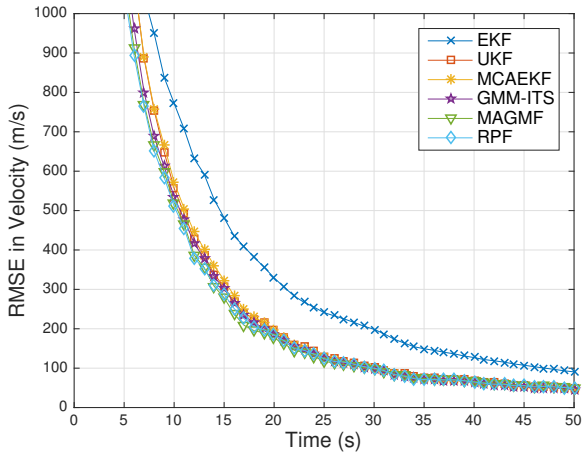

Figure 127: Total Velocity RMSE for Scenario Z3 ($C_B = 3$, $R_B = 1$)

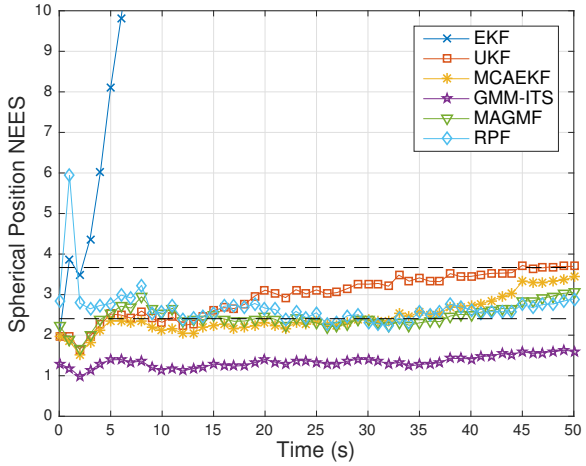

Figure 128: Spherical Position NEES for Scenario Z3 ($C_B = 3$, $R_B = 1$)

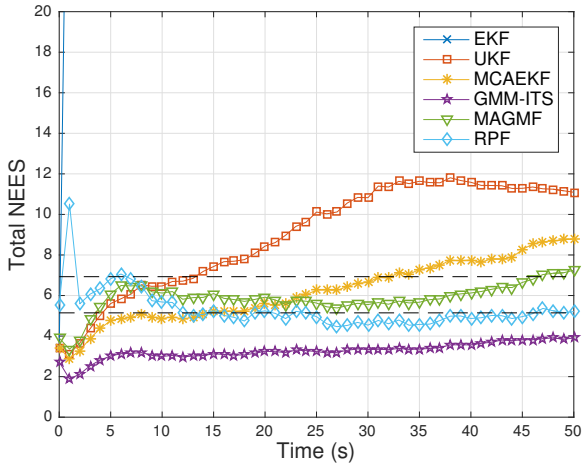

Figure 129: Total NEES for Scenario Z3 ($C_B = 3$, $R_B = 1$)

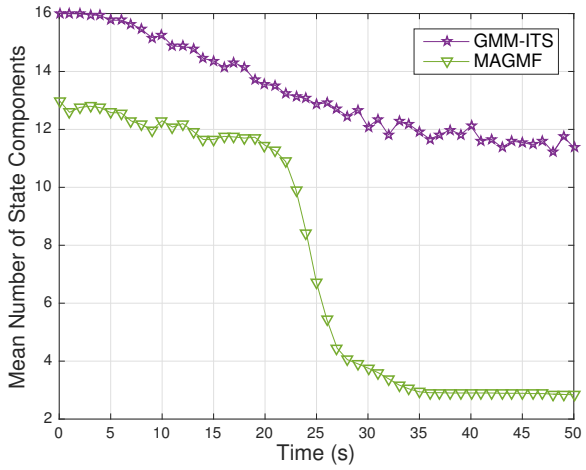

Figure 130: Number of State Components for Scenario Z3 ($C_B = 3$, $R_B = 1$)

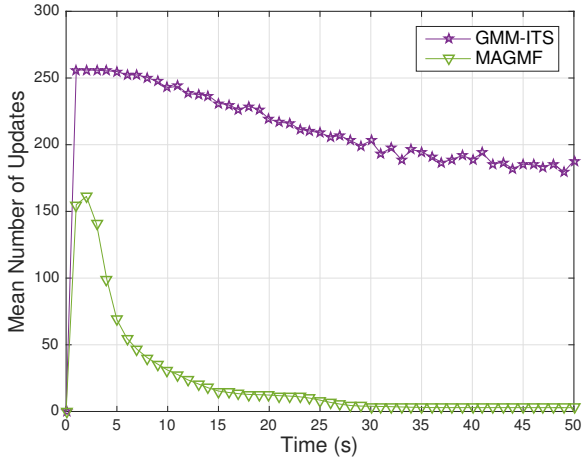

Figure 131: Number of Kalman Updates for Scenario Z3 ($C_B = 3$, $R_B = 1$)

4.8 Conclusions

In this chapter, the adaptive Gaussian mixture filter process presented for 2D monostatic radar measurements in Chapter 2 was extended to 3D measurements in range, azimuth, and elevation. Equivalent expressions for bias significance and KL divergence of single-Gaussian measurement models were derived for the 3D case. These KL divergence expressions were as the basis of numerical fitting experiments using the EM algorithm to fit Gaussian mixtures of various sizes to 3D radar measurements.

Based on the results of these experiments it was found that, unlike the 2D case, the KL divergence that can be achieved by a GM fit to a 3D measurement depends on the total bias significance (C_B) and the ratio of azimuth and elevation bias significances (R_B). Focusing on the case of $R_B = 1$, the geometric arrangements of many of these mixture solutions were characterized and cataloged. Where multiple geometries exist for a given number of Gaussian components (N_G), certain geometries were identified as preferable based on the properties of regularity, KL divergence performance, and size of angular support.

In a similar manner to that in Chapter 2, a lookup table for the weight, means, and covariances for GM solutions derived using the EM algorithm was generated by appropriate transformation of the results. In this chapter, this lookup table was restricted to cover the case of unity bias significance ratio ($R_B = 1$), and two procedures for adapting the results of this table (suboptimally) to other bias significance ratios were presented and shown to give reasonable results. This lookup table was used to build a 3D version of the MAGMF filter originally presented in Chapter 2.

The performance of the 3D MAGMF was compared to existing nonlinear estimation techniques for the contact-lens problem in 3D. These included the MAGMF, RPF and GMM-ITS. Figure 132 shows the relative performance of these filters in the most difficult filtering scenarios from this work. The MAGMF was found to have excellent range estimation performance, excellent overall estimation performance in position

and velocity, and good consistency in both Cartesian and spherical space. Additionally, due to the adaptation of the number of GMs based on the KL divergence of the distributions and the conditioning approach used to locate the measurement components in the vicinity of the state angle support, the MAGMF requires significantly fewer computations than the GMM-ITS and RPF while providing similar or superior performance. Much of the contents of this chapter are represented in a submitted journal paper [11].

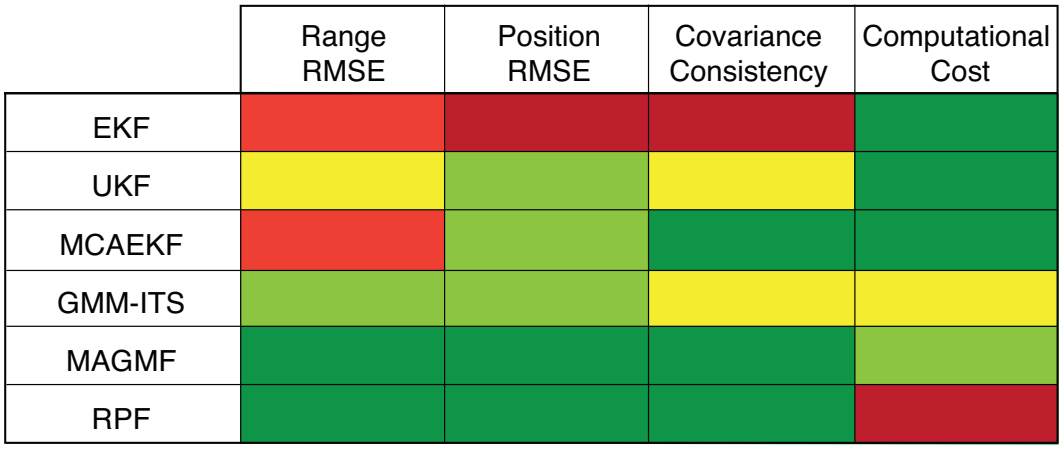

Figure 132: Performance Summary of Filters in 3D Monostatic Simulations

CHAPTER V

CONCLUSIONS AND FUTURE RESEARCH

This dissertation provides several key contributions to the state-of-the art in tracking with measurements that suffer from the contact-lens effect. This section provides an overview of the accomplishments, their implications for tracker design, and suggested future research to take advantage of the demonstrated improvements.

Chapter 2 concentrates on the contact-lens problem in the well-studied 2D monostatic case. Previous research focused on adaptation of single-Gaussian solutions to provide consistent filtering results when tracking with measurements suffering from the contact-lens effect. However, these solutions result in degraded range estimation performance and inflate the measurement-to-track gating region.

The first major contribution in this chapter is the derivation of the KL divergence of single-Gaussian fits to the polar measurement distribution from the true distribution in terms of the previously published metric of bias significance. KL divergence is a widely used information-theoretic quantity which is directly related to estimator efficiency and maximum likelihood fitting. The increased KL divergence of the single-Gaussian fits from the true distribution is interpreted geometrically and linked to the increased area of the gating region.

The second major contribution in Chapter 2 is the discovery that measurements with the same bias significance have the same KL divergence when represented by an ML-optimal GM with specified number of components. This fact is combined with a novel normalization procedure to allow the ML-optimal GM parameters of weight, mean, and covariance to be stored in a lookup table of reasonable size. This allows ML GM models of contact-lens measurements to be generated in real-time and used in

a filtering algorithm with low computational requirements. Previous work using GMs to address the contact-lens problem did not make use of ML-optimal parameters and relied on ad-hoc methods for deriving the GMs. The use of ML-optimal parameters allows a smaller number of mixture components to be used for the measurement model than an ad-hoc approach.

An additional contribution related to the ML-optimal lookup table is that the KL divergence of each ML GM fit is stored in the table based on the bias significance of the input measurement. This allows the GM model size to be adaptively selected to limit the KL divergence of the model to a desired value. Previous published approaches did not specify a rigorous way to choose the mixture size or adapt it during tracking.

The third major contribution of Chapter 2 addresses the fact that the state and measurement PDFs in a filter for the contact-lens problem will often have mismatched support once the filter has converged. If a naïve approach is used to model the mixture PDF with a GM, a large number of components is needed, as the model must have good fidelity over the whole support of the state PDF. However, most of these components will have negligible weight after the filter update. Chapter 2 presents a method that can be used to pre-condition the measurement to the state support in angle. The size of the GM model needed for this conditioned measurement is much smaller that the naïve approach. Additionally, the size of the model converges rapidly with the size of the state uncertainty. After the mixture filter update occurs, the conditioning factor is removed by an inverse filtering process.

These contributions are leveraged in Chapter 2 to form a new GM filter, the MAGMF. This filter is run in comparison to other filtering techniques which have been proposed to address the 2D contact-lens problem. The MAGMF is able to provide excellent range RMSE and good filter consistency while using fewer computations than a particle filter and fewer GM components than previously proposed

GM approaches. Additionally, the gating region size of the MAGMF is greatly reduced compared to single-Gaussian filters while still retaining the ability to gate true measurements with the desired probability of gating.

Chapter 3 extends the MAGMF developed in Chapter 2 to 2D bistatic radar. The first major contribution is the extension of the definition of bias significance originally developed in a monostatic context to bistatic radar. A concise formula is given which allows evaluation of the KL divergence of a single-Gaussian bistatic measurement model from the true measurement distribution using this bistatic bias significance.

The second major contribution of Chapter 3 is the application of the same ML modeling method for GM parameters used in Chapter 2 to bistatic radar measurements. A key discovery shows that the same lookup tables used to generate monostatic GM models can be used to generate bistatic GM models with a slight modification of the normalization procedure. The MAGMF filter from Chapter 2 was modified to use these bistatic measurement models and was run for comparison with other methods proposed to perform bistatic tracking. The MAGMF provides excellent RMSE performance in bistatic range and overall better position and velocity RMSE than previously studied single-Gaussian approaches while requiring significantly fewer computations than a particle filter.

Chapter 4 applies the same techniques to 3D monostatic radar measurements (RAE). The first major contribution of this chapter is the extension of the bias significance definition to 3D radar, which allows a similar evaluation of the KL divergence of a single-Gaussian fit to the true measurement. The second contribution is the characterization of the KL divergence performance and geometric layout of ML GM fits to 3D radar measurements. Unlike the 2D monostatic and bistatic cases, the KL divergence of the ML models is not determined by bias significance alone. The ratio between azimuth and elevation error in the measurement also affects the KL divergence, with the highest divergence occurring when they are equal. This fact was

used to define a new measurement parameter, the bias significance ratio. A labeling method was created to describe the geometric layout of the components in the GM solutions, and favorable layouts were identified for a given number of components.

The third contribution of Chapter 4 is a normalization technique that can be applied to 3D ML GM results to allow them to be stored in a lookup table. This lookup table is currently restricted to measurements with unity bias significance ratio and particular GM solution geometries identified to be favorable for use. By exploiting the circular symmetry of the ML solutions a compact representation of the GM parameters is possible. Examples of these parameters are presented in Appendix B.

The fourth contribution of Chapter 4 is the identification of two possible methods to sub-optimally extend the use of the ML lookup table to measurements with non-unity bias significance ratio. The KL divergence of these methods are studied in comparison to the optimal solutions to allow their modeling performance to be validated for use in a filtering algorithm.

Finally, the MAGMF filter of Chapter 2 was modified to operate using the 3D ML lookup table. The performance of this filter is compared to other filters proposed to address the contact-lens problem in 3D. The 3D MAGMF demonstrated excellent range RMSE performance and covariance consistency while having a low computational cost compared to the alternative GM approaches and particle filters. Additionally, unlike previous approaches to the contact-lens problem, the computational cost of the MAGMF decreases over time as the state covariance shrinks.

This dissertation demonstrates the effectiveness of an ML GM modeling approach in conjunction with pre-conditioning of the measurements to better match the prior state support in solving the contact-lens problem in radar tracking. The size of the gating region of the MAGMF was shown to be significantly smaller than previously proposed filtering schemes and provides increased robustness against false alarms in the association process. However, the performance of the MAGMF has not been

studied in a multi-sensor, multi-target scenario. Future research should focus on application of the GM tracks produced by the MAGMF to a sensor fusion system with track-to-track correlation between sensors tracking multiple targets. Since the GM tracks preserve range accuracy with higher fidelity than their single-Gaussian counterparts, correlation of multi-target tracks between sensors should be become possible for more closely spaced targets. However, these advantages in sensor fusion would require the traditional sensor track reporting format of a single mean and covariance matrix to be modified to support GMs.

In Chapter 4, an ad-hoc method of extending the lookup table results to measurements with non-unity bias significance ratio was presented. This method yielded reasonable results when applied in the MAGMF. However, further research should derive new lookup tables and parameter normalization techniques to allow these measurements to be more accurately represented using a GM. This would further reduce the computational cost of the MAGMF, as the optimal GM models for these measurements can achieve a lower KL divergence for a given number of components.

A further potential use of the MAGMF is long-range tracking of maneuvering targets with multiple dynamics models. In this case, improved range accuracy is essential for reliable detection of maneuver accelerations, which drives the correct selection of dynamics models. Future research should investigate the application of the MAGMF in a multiple-model filtering scheme.

APPENDIX A

ML LOOKUP TABLES FOR 2D MONOSTATIC AND BISTATIC GM FITS

This appendix lists the tables necessary to generate ML optimal GM fits for two-dimensional monostatic and bistatic measurements. The component indices are specified as positive or negative numbers based on the normalized angular displacement of the component mean $\Delta\hat{\theta}_k$ relative to the overall distribution mean. For example, $k = 1$ will have the lowest positive angular displacement and $k = -1$ has the greatest negative angular displacement. In the case that an odd number of components is used, a $k = 0$ component will be present centered on the mean. This component has a normalized angular displacement of $\Delta\hat{\theta}_0 = 0$ and correlation coefficent $\rho_{r\theta,0} = 0$ for all GM fits due to the symmetry of the distribution. Therefore, these parameters are omitted from the tables to save space. For bistatic measurements, the parameters $\Delta\hat{r}_k$, $\Delta\hat{\theta}_k$, $\hat{\sigma}_{r,k}$, and $\rho_{r\theta,k}$ in these tables give the values of $\Delta\hat{R}_{S,k}$, $\Delta\hat{\theta}_{R,k}$, $\hat{\sigma}_{S,k}$, and $\rho_{S\theta,k}$, respectively.

Table 9: KL Divergence and Component $k = \pm 1$ Parameters for $N_G = 2$

C_B	KL	$\omega_{\pm 1}$	$\Delta\hat{r}_{\pm 1}$	$\Delta\hat{\theta}_{\pm 1}$	$\hat{\sigma}_{r,\pm 1}$	$\hat{\sigma}_{\theta,\pm 1}$	$\rho_{r\theta,\pm 1}$
0.20	0.004	0.500	-0.75	± 0.50	0.95	0.87	∓ 0.21
0.24	0.006	0.500	-0.73	± 0.52	0.94	0.85	∓ 0.22
0.29	0.010	0.500	-0.70	± 0.55	0.92	0.84	∓ 0.23
0.34	0.016	0.500	-0.68	± 0.57	0.89	0.82	∓ 0.24
0.41	0.024	0.500	-0.65	± 0.59	0.86	0.81	∓ 0.25
0.49	0.038	0.500	-0.62	± 0.62	0.82	0.79	∓ 0.26
0.59	0.057	0.500	-0.59	± 0.64	0.78	0.77	∓ 0.27
0.71	0.084	0.500	-0.55	± 0.67	0.74	0.74	∓ 0.27
0.85	0.121	0.500	-0.51	± 0.70	0.70	0.71	∓ 0.28
1.01	0.169	0.500	-0.48	± 0.72	0.66	0.69	∓ 0.30
1.21	0.229	0.500	-0.45	± 0.74	0.62	0.67	∓ 0.32
1.46	0.300	0.500	-0.42	± 0.76	0.59	0.65	∓ 0.35
1.74	0.386	0.500	-0.40	± 0.77	0.56	0.64	∓ 0.39
2.09	0.483	0.500	-0.39	± 0.78	0.53	0.63	∓ 0.43
2.50	0.594	0.500	-0.38	± 0.78	0.51	0.62	∓ 0.46

Table 10: KL Divergence and Component $k = 0$ Parameters for $N_G = 3$

C_B	KL	ω_0	$\Delta\hat{r}_0$	$\hat{\sigma}_{r,0}$	$\hat{\sigma}_{\theta,0}$
0.50	0.009	0.571	-0.47	0.77	0.68
0.57	0.012	0.570	-0.44	0.74	0.66
0.65	0.017	0.566	-0.41	0.69	0.64
0.73	0.024	0.562	-0.38	0.65	0.62
0.83	0.032	0.557	-0.36	0.60	0.60
0.95	0.044	0.553	-0.34	0.56	0.58
1.08	0.058	0.547	-0.32	0.51	0.56
1.22	0.075	0.543	-0.29	0.47	0.54
1.39	0.096	0.539	-0.27	0.43	0.52
1.58	0.121	0.536	-0.25	0.39	0.50
1.80	0.152	0.533	-0.23	0.36	0.48
2.04	0.187	0.532	-0.22	0.33	0.47
2.32	0.227	0.532	-0.20	0.30	0.46
2.64	0.272	0.531	-0.19	0.27	0.44
3.00	0.321	0.529	-0.18	0.25	0.43

Table 11: Component $k = \pm 1$ Parameters for $N_G = 3$

C_B	$\omega_{\pm 1}$	$\Delta\hat{r}_{\pm 1}$	$\Delta\hat{\theta}_{\pm 1}$	$\hat{\sigma}_{r,\pm 1}$	$\hat{\sigma}_{\theta,\pm 1}$	$\rho_{r\theta,\pm 1}$
0.50	0.214	-0.49	± 1.11	0.81	0.71	∓ 0.31
0.57	0.215	-0.47	± 1.13	0.78	0.69	∓ 0.31
0.65	0.217	-0.45	± 1.15	0.75	0.67	∓ 0.31
0.73	0.219	-0.43	± 1.17	0.71	0.66	∓ 0.31
0.83	0.222	-0.41	± 1.18	0.67	0.64	∓ 0.32
0.95	0.224	-0.38	± 1.20	0.64	0.62	∓ 0.32
1.08	0.226	-0.36	± 1.21	0.60	0.60	∓ 0.33
1.22	0.228	-0.34	± 1.22	0.56	0.59	∓ 0.34
1.39	0.230	-0.33	± 1.23	0.53	0.58	∓ 0.35
1.58	0.232	-0.31	± 1.24	0.50	0.56	∓ 0.37
1.80	0.233	-0.30	± 1.25	0.47	0.55	∓ 0.38
2.04	0.234	-0.29	± 1.26	0.45	0.54	∓ 0.40
2.32	0.234	-0.28	± 1.27	0.43	0.53	∓ 0.42
2.64	0.235	-0.27	± 1.28	0.41	0.52	∓ 0.44
3.00	0.235	-0.27	± 1.28	0.39	0.51	∓ 0.46

Table 12: KL Divergence and Component $k = \pm 1$ Parameters for $N_G = 4$

C_B	KL	$\omega_{\pm 1}$	$\Delta\hat{r}_{\pm 1}$	$\Delta\hat{\theta}_{\pm 1}$	$\hat{\sigma}_{r,\pm 1}$	$\hat{\sigma}_{\theta,\pm 1}$	$\rho_{r\theta,\pm 1}$
1.00	0.016	0.396	-0.27	± 0.50	0.54	0.52	∓ 0.08
1.22	0.027	0.392	-0.24	± 0.51	0.47	0.49	∓ 0.08
1.50	0.042	0.388	-0.22	± 0.50	0.40	0.46	∓ 0.08
1.63	0.051	0.387	-0.21	± 0.50	0.38	0.45	∓ 0.08
1.78	0.061	0.385	-0.20	± 0.50	0.35	0.44	∓ 0.08
1.94	0.072	0.384	-0.19	± 0.50	0.33	0.43	∓ 0.08
2.12	0.085	0.382	-0.18	± 0.50	0.30	0.42	∓ 0.08
2.31	0.099	0.381	-0.17	± 0.50	0.28	0.41	∓ 0.08
2.51	0.115	0.379	-0.16	± 0.50	0.26	0.40	∓ 0.07
2.74	0.133	0.379	-0.15	± 0.50	0.25	0.39	∓ 0.07
2.98	0.152	0.377	-0.15	± 0.50	0.23	0.38	∓ 0.07
3.25	0.174	0.376	-0.14	± 0.50	0.21	0.38	∓ 0.06
3.54	0.197	0.375	-0.13	± 0.50	0.20	0.37	∓ 0.06
3.86	0.223	0.373	-0.13	± 0.50	0.19	0.36	∓ 0.06
4.21	0.250	0.372	-0.12	± 0.50	0.17	0.35	∓ 0.06
4.59	0.280	0.371	-0.12	± 0.50	0.16	0.35	∓ 0.06
5.00	0.310	0.369	-0.11	± 0.50	0.15	0.34	∓ 0.06

Table 13: Component $k = \pm 2$ Parameters for $N_G = 4$

C_B	$\omega_{\pm 2}$	$\Delta\hat{r}_{\pm 2}$	$\Delta\hat{\theta}_{\pm 2}$	$\hat{\sigma}_{r,\pm 2}$	$\hat{\sigma}_{\theta,\pm 2}$	$\rho_{r\theta,\pm 2}$
1.00	0.104	-0.32	± 1.57	0.60	0.57	∓ 0.34
1.22	0.108	-0.29	± 1.59	0.54	0.55	∓ 0.35
1.50	0.112	-0.27	± 1.60	0.48	0.52	∓ 0.36
1.63	0.113	-0.26	± 1.61	0.46	0.52	∓ 0.37
1.78	0.115	-0.25	± 1.61	0.44	0.51	∓ 0.38
1.94	0.116	-0.25	± 1.61	0.42	0.50	∓ 0.39
2.12	0.118	-0.24	± 1.61	0.40	0.49	∓ 0.40
2.31	0.119	-0.23	± 1.62	0.39	0.48	∓ 0.41
2.51	0.121	-0.23	± 1.62	0.37	0.48	∓ 0.42
2.74	0.121	-0.23	± 1.62	0.36	0.47	∓ 0.43
2.98	0.123	-0.22	± 1.62	0.35	0.47	∓ 0.44
3.25	0.124	-0.22	± 1.62	0.34	0.46	∓ 0.46
3.54	0.125	-0.21	± 1.62	0.33	0.46	∓ 0.47
3.86	0.127	-0.21	± 1.62	0.32	0.45	∓ 0.49
4.21	0.128	-0.21	± 1.62	0.31	0.45	∓ 0.50
4.59	0.129	-0.21	± 1.62	0.31	0.45	∓ 0.52
5.00	0.131	-0.21	± 1.61	0.30	0.45	∓ 0.53

Table 14: KL Divergence and Component $k = 0$ Parameters for $N_G = 5$

C_B	KL	ω_0	$\Delta\hat{r}_0$	$\hat{\sigma}_{r,0}$	$\hat{\sigma}_{\theta,0}$
1.60	0.021	0.377	−0.18	0.38	0.42
1.80	0.027	0.373	−0.16	0.35	0.40
2.00	0.034	0.370	−0.15	0.31	0.39
2.21	0.042	0.368	−0.15	0.29	0.38
2.44	0.052	0.364	−0.14	0.26	0.37
2.69	0.063	0.360	−0.13	0.24	0.36
2.97	0.076	0.357	−0.12	0.22	0.35
3.28	0.090	0.353	−0.12	0.20	0.34
3.62	0.107	0.349	−0.11	0.18	0.33
4.00	0.125	0.345	−0.10	0.17	0.32
4.42	0.146	0.341	−0.10	0.15	0.31
4.88	0.170	0.338	−0.09	0.14	0.30
5.38	0.195	0.335	−0.08	0.13	0.29
5.94	0.224	0.332	−0.08	0.12	0.28
6.56	0.253	0.329	−0.08	0.11	0.28
7.25	0.286	0.326	−0.07	0.10	0.27
8.00	0.322	0.324	−0.07	0.09	0.26

Table 15: Component $k = \pm 1$ Parameters for $N_G = 5$

C_B	$\omega_{\pm 1}$	$\Delta\hat{r}_{\pm 1}$	$\Delta\hat{\theta}_{\pm 1}$	$\hat{\sigma}_{r,\pm 1}$	$\hat{\sigma}_{\theta,\pm 1}$	$\rho_{r\theta,\pm 1}$
1.60	0.254	-0.19	± 0.89	0.38	0.43	∓ 0.12
1.80	0.254	-0.17	± 0.88	0.35	0.42	∓ 0.12
2.00	0.254	-0.16	± 0.88	0.32	0.40	∓ 0.12
2.21	0.254	-0.15	± 0.88	0.29	0.39	∓ 0.11
2.44	0.254	-0.15	± 0.88	0.27	0.38	∓ 0.11
2.69	0.254	-0.14	± 0.87	0.24	0.37	∓ 0.11
2.97	0.254	-0.13	± 0.87	0.22	0.36	∓ 0.11
3.28	0.255	-0.12	± 0.87	0.21	0.35	∓ 0.10
3.62	0.255	-0.12	± 0.86	0.19	0.34	∓ 0.10
4.00	0.255	-0.11	± 0.85	0.17	0.33	∓ 0.10
4.42	0.256	-0.10	± 0.85	0.16	0.33	∓ 0.09
4.88	0.256	-0.10	± 0.85	0.15	0.32	∓ 0.09
5.38	0.256	-0.10	± 0.84	0.14	0.31	∓ 0.09
5.94	0.255	-0.09	± 0.84	0.13	0.31	∓ 0.09
6.56	0.255	-0.09	± 0.83	0.12	0.30	∓ 0.09
7.25	0.254	-0.08	± 0.83	0.11	0.29	∓ 0.10
8.00	0.253	-0.08	± 0.82	0.10	0.29	∓ 0.10

Table 16: Component $k = \pm 2$ Parameters for $N_G = 5$

C_B	$\omega_{\pm 2}$	$\Delta\hat{r}_{\pm 2}$	$\Delta\hat{\theta}_{\pm 2}$	$\hat{\sigma}_{r,\pm 2}$	$\hat{\sigma}_{\theta,\pm 2}$	$\rho_{r\theta,\pm 2}$
1.60	0.058	-0.23	± 1.89	0.45	0.49	∓ 0.37
1.80	0.060	-0.22	± 1.90	0.42	0.48	∓ 0.38
2.00	0.061	-0.21	± 1.89	0.40	0.47	∓ 0.39
2.21	0.063	-0.21	± 1.89	0.37	0.46	∓ 0.40
2.44	0.064	-0.20	± 1.89	0.36	0.45	∓ 0.41
2.69	0.066	-0.20	± 1.89	0.34	0.44	∓ 0.42
2.97	0.067	-0.19	± 1.89	0.32	0.44	∓ 0.43
3.28	0.069	-0.19	± 1.89	0.31	0.43	∓ 0.45
3.62	0.070	-0.19	± 1.88	0.30	0.43	∓ 0.46
4.00	0.072	-0.18	± 1.88	0.29	0.42	∓ 0.48
4.42	0.074	-0.18	± 1.87	0.28	0.42	∓ 0.49
4.88	0.075	-0.18	± 1.87	0.28	0.41	∓ 0.51
5.38	0.077	-0.18	± 1.86	0.27	0.41	∓ 0.53
5.94	0.079	-0.18	± 1.85	0.27	0.41	∓ 0.54
6.56	0.081	-0.17	± 1.84	0.26	0.41	∓ 0.56
7.25	0.082	-0.17	± 1.84	0.26	0.41	∓ 0.57
8.00	0.085	-0.17	± 1.83	0.26	0.41	∓ 0.58

Table 17: KL Divergence and Component $k = \pm 1$ Parameters for $N_G = 6$

C_B	KL	$\omega_{\pm 1}$	$\Delta\hat{r}_{\pm 1}$	$\Delta\hat{\theta}_{\pm 1}$	$\hat{\sigma}_{r,\pm 1}$	$\hat{\sigma}_{\theta,\pm 1}$	$\rho_{r\theta,\pm 1}$
2.20	0.020	0.300	-0.13	± 0.38	0.29	0.36	∓ 0.04
2.40	0.025	0.299	-0.12	± 0.38	0.27	0.35	∓ 0.04
2.60	0.029	0.296	-0.12	± 0.38	0.25	0.34	∓ 0.04
2.80	0.035	0.294	-0.11	± 0.38	0.23	0.33	∓ 0.04
3.00	0.040	0.293	-0.11	± 0.37	0.22	0.33	∓ 0.04
3.31	0.049	0.291	-0.10	± 0.37	0.20	0.32	∓ 0.04
3.66	0.059	0.288	-0.10	± 0.37	0.18	0.31	∓ 0.04
4.04	0.071	0.285	-0.09	± 0.37	0.16	0.30	∓ 0.04
4.46	0.084	0.282	-0.08	± 0.36	0.15	0.29	∓ 0.04
4.92	0.100	0.279	-0.08	± 0.36	0.14	0.28	∓ 0.04
5.43	0.117	0.276	-0.07	± 0.36	0.13	0.27	∓ 0.04
6.00	0.136	0.274	-0.07	± 0.35	0.11	0.26	∓ 0.03
6.62	0.158	0.271	-0.07	± 0.35	0.10	0.26	∓ 0.03
7.31	0.182	0.269	-0.06	± 0.35	0.10	0.25	∓ 0.03
8.08	0.207	0.266	-0.06	± 0.34	0.09	0.24	∓ 0.03
8.92	0.235	0.264	-0.06	± 0.34	0.08	0.24	∓ 0.03
9.84	0.266	0.262	-0.05	± 0.34	0.08	0.23	∓ 0.03
10.87	0.299	0.260	-0.05	± 0.34	0.07	0.22	∓ 0.03
12.00	0.334	0.258	-0.05	± 0.33	0.06	0.22	∓ 0.03

Table 18: Component $k = \pm 2$ Parameters for $N_G = 6$

C_B	$\omega_{\pm 2}$	$\Delta\hat{r}_{\pm 2}$	$\Delta\hat{\theta}_{\pm 2}$	$\hat{\sigma}_{r,\pm 2}$	$\hat{\sigma}_{\theta,\pm 2}$	$\rho_{r\theta,\pm 2}$
2.20	0.165	-0.14	± 1.18	0.29	0.38	∓ 0.14
2.40	0.166	-0.13	± 1.18	0.27	0.37	∓ 0.14
2.60	0.167	-0.13	± 1.17	0.25	0.36	∓ 0.13
2.80	0.168	-0.12	± 1.17	0.23	0.35	∓ 0.13
3.00	0.169	-0.12	± 1.16	0.22	0.34	∓ 0.13
3.31	0.170	-0.11	± 1.16	0.20	0.34	∓ 0.13
3.66	0.171	-0.11	± 1.15	0.18	0.33	∓ 0.13
4.04	0.173	-0.10	± 1.14	0.17	0.32	∓ 0.12
4.46	0.174	-0.09	± 1.13	0.16	0.31	∓ 0.12
4.92	0.175	-0.09	± 1.13	0.14	0.30	∓ 0.12
5.43	0.177	-0.09	± 1.12	0.13	0.30	∓ 0.11
6.00	0.178	-0.08	± 1.11	0.12	0.29	∓ 0.11
6.62	0.179	-0.08	± 1.10	0.11	0.28	∓ 0.11
7.31	0.180	-0.07	± 1.09	0.10	0.28	∓ 0.11
8.08	0.180	-0.07	± 1.09	0.10	0.27	∓ 0.11
8.92	0.181	-0.07	± 1.08	0.09	0.26	∓ 0.12
9.84	0.181	-0.07	± 1.07	0.08	0.26	∓ 0.12
10.87	0.181	-0.06	± 1.06	0.08	0.25	∓ 0.13
12.00	0.181	-0.06	± 1.05	0.07	0.25	∓ 0.13

Table 19: Component $k = \pm 3$ Parameters for $N_G = 6$

C_B	$\omega_{\pm 3}$	$\Delta\hat{r}_{\pm 3}$	$\Delta\hat{\theta}_{\pm 3}$	$\hat{\sigma}_{r,\pm 3}$	$\hat{\sigma}_{\theta,\pm 3}$	$\rho_{r\theta,\pm 3}$
2.20	0.035	-0.19	± 2.13	0.36	0.44	∓ 0.39
2.40	0.036	-0.18	± 2.13	0.34	0.43	∓ 0.40
2.60	0.037	-0.18	± 2.12	0.33	0.43	∓ 0.41
2.80	0.038	-0.17	± 2.12	0.32	0.42	∓ 0.42
3.00	0.039	-0.17	± 2.12	0.31	0.42	∓ 0.43
3.31	0.040	-0.17	± 2.11	0.29	0.41	∓ 0.44
3.66	0.041	-0.17	± 2.10	0.28	0.40	∓ 0.45
4.04	0.042	-0.16	± 2.10	0.27	0.40	∓ 0.47
4.46	0.044	-0.16	± 2.09	0.26	0.40	∓ 0.48
4.92	0.045	-0.16	± 2.08	0.26	0.39	∓ 0.50
5.43	0.047	-0.16	± 2.07	0.25	0.39	∓ 0.52
6.00	0.048	-0.16	± 2.06	0.24	0.39	∓ 0.53
6.62	0.050	-0.16	± 2.05	0.24	0.39	∓ 0.55
7.31	0.052	-0.16	± 2.04	0.24	0.39	∓ 0.56
8.08	0.054	-0.15	± 2.03	0.23	0.39	∓ 0.58
8.92	0.055	-0.15	± 2.01	0.23	0.39	∓ 0.59
9.84	0.057	-0.15	± 2.00	0.23	0.39	∓ 0.60
10.87	0.059	-0.15	± 1.99	0.23	0.39	∓ 0.61
12.00	0.061	-0.16	± 1.97	0.23	0.39	∓ 0.62

Table 20: KL Divergence and Component $k = 0$ Parameters for $N_G = 7$

C_B	KL	ω_0	$\Delta\hat{r}_0$	$\hat{\sigma}_{r,0}$	$\hat{\sigma}_{\theta,0}$
4.00	0.040	0.270	−0.08	0.16	0.28
4.42	0.048	0.267	−0.07	0.15	0.27
4.88	0.059	0.264	−0.07	0.14	0.26
5.38	0.070	0.260	−0.07	0.12	0.26
5.94	0.083	0.257	−0.06	0.11	0.25
6.56	0.097	0.255	−0.06	0.10	0.24
7.25	0.114	0.251	−0.05	0.09	0.23
8.00	0.133	0.248	−0.05	0.09	0.23
8.83	0.153	0.244	−0.05	0.08	0.22
9.75	0.176	0.242	−0.05	0.07	0.21
10.77	0.201	0.239	−0.04	0.07	0.21
11.89	0.227	0.236	−0.04	0.06	0.20
13.13	0.256	0.234	−0.04	0.06	0.19
14.49	0.287	0.232	−0.04	0.05	0.19
16.00	0.322	0.230	−0.03	0.05	0.19

Table 21: Component $k = \pm 1$ Parameters for $N_G = 7$

C_B	$\omega_{\pm 1}$	$\Delta\hat{r}_{\pm 1}$	$\Delta\hat{\theta}_{\pm 1}$	$\hat{\sigma}_{r,\pm 1}$	$\hat{\sigma}_{\theta,\pm 1}$	$\rho_{r\theta,\pm 1}$
4.00	0.223	−0.08	±0.66	0.16	0.29	∓0.06
4.42	0.221	−0.08	±0.66	0.15	0.28	∓0.06
4.88	0.220	−0.07	±0.65	0.14	0.27	∓0.06
5.38	0.218	−0.07	±0.64	0.12	0.26	∓0.06
5.94	0.217	−0.06	±0.64	0.11	0.25	∓0.06
6.56	0.216	−0.06	±0.63	0.10	0.25	∓0.06
7.25	0.214	−0.06	±0.62	0.09	0.24	∓0.05
8.00	0.213	−0.05	±0.62	0.09	0.23	∓0.05
8.83	0.212	−0.05	±0.61	0.08	0.22	∓0.05
9.75	0.210	−0.05	±0.61	0.07	0.22	∓0.05
10.77	0.209	−0.05	±0.60	0.07	0.21	∓0.05
11.89	0.208	−0.04	±0.60	0.06	0.21	∓0.05
13.13	0.207	−0.04	±0.59	0.06	0.20	∓0.05
14.49	0.205	−0.04	±0.58	0.05	0.20	∓0.05
16.00	0.204	−0.04	±0.58	0.05	0.19	∓0.05

Table 22: Component $k = \pm 2$ Parameters for $N_G = 7$

C_B	$\omega_{\pm 2}$	$\Delta\hat{r}_{\pm 2}$	$\Delta\hat{\theta}_{\pm 2}$	$\hat{\sigma}_{r,\pm 2}$	$\hat{\sigma}_{\theta,\pm 2}$	$\rho_{r\theta,\pm 2}$
4.00	0.117	-0.09	± 1.38	0.17	0.31	∓ 0.14
4.42	0.118	-0.09	± 1.37	0.15	0.30	∓ 0.14
4.88	0.120	-0.08	± 1.36	0.14	0.29	∓ 0.13
5.38	0.122	-0.08	± 1.35	0.13	0.29	∓ 0.13
5.94	0.124	-0.08	± 1.34	0.12	0.28	∓ 0.13
6.56	0.125	-0.07	± 1.33	0.11	0.27	∓ 0.13
7.25	0.127	-0.07	± 1.32	0.10	0.27	∓ 0.13
8.00	0.129	-0.07	± 1.31	0.09	0.26	∓ 0.13
8.83	0.130	-0.06	± 1.30	0.09	0.25	∓ 0.13
9.75	0.131	-0.06	± 1.29	0.08	0.25	∓ 0.13
10.77	0.132	-0.06	± 1.28	0.07	0.24	∓ 0.13
11.89	0.133	-0.06	± 1.26	0.07	0.24	∓ 0.14
13.13	0.134	-0.05	± 1.25	0.06	0.23	∓ 0.14
14.49	0.134	-0.05	± 1.24	0.06	0.23	∓ 0.15
16.00	0.135	-0.05	± 1.23	0.05	0.23	∓ 0.16

Table 23: Component $k = \pm 3$ Parameters for $N_G = 7$

C_B	$\omega_{\pm 3}$	$\Delta\hat{r}_{\pm 3}$	$\Delta\hat{\theta}_{\pm 3}$	$\hat{\sigma}_{r,\pm 3}$	$\hat{\sigma}_{\theta,\pm 3}$	$\rho_{r\theta,\pm 3}$
4.00	0.026	-0.15	± 2.29	0.26	0.38	∓ 0.46
4.42	0.027	-0.15	± 2.28	0.25	0.38	∓ 0.47
4.88	0.028	-0.15	± 2.27	0.24	0.38	∓ 0.49
5.38	0.029	-0.14	± 2.26	0.23	0.37	∓ 0.50
5.94	0.031	-0.14	± 2.24	0.23	0.37	∓ 0.52
6.56	0.032	-0.14	± 2.23	0.22	0.37	∓ 0.54
7.25	0.033	-0.14	± 2.22	0.22	0.37	∓ 0.55
8.00	0.035	-0.14	± 2.20	0.22	0.37	∓ 0.57
8.83	0.036	-0.14	± 2.19	0.22	0.37	∓ 0.58
9.75	0.038	-0.14	± 2.17	0.21	0.37	∓ 0.59
10.77	0.039	-0.14	± 2.16	0.21	0.37	∓ 0.60
11.89	0.041	-0.14	± 2.14	0.21	0.37	∓ 0.61
13.13	0.043	-0.14	± 2.12	0.21	0.37	∓ 0.62
14.49	0.044	-0.14	± 2.11	0.21	0.37	∓ 0.63
16.00	0.046	-0.14	± 2.09	0.21	0.37	∓ 0.63

APPENDIX B

ML LOOKUP TABLES FOR 3D MONOSTATIC GM FITS WITH BIAS SIGNIFICANCE RATIOS OF ONE

This appendix lists the tables necessary to generate ML optimal GM fits for three-dimensional monostatic measurements with bias significance ratio of one. The GM models represented are identified by their geometry code as described in Section 4.3. In these tables, the index expression $(\cdot)_c$ indicates the parameters associated with a central component, while $(\cdot)_\ell$ indicates the parameters associated with ring ℓ. As the angle orientation values $\psi_{\ell,k}$ do not change with bias significance, they are listed in a separate table for all geometries. The central component parameters $\Delta\hat{\alpha}_c$ and $\Delta\hat{\varepsilon}_c$ are zero and not listed in the tables. Also, all cross-correlation values for the central component are defined to be zero, and the normalized azimuth and elevation variances are identical and expressed by a single parameter $\hat{\sigma}_{\alpha\varepsilon,c}$. For the ring components, only the range-to-angle displacement correlation is non-zero and given by $\rho_{r\hat{\theta},\ell}$. The other correlation values are assumed to be zero and not listed in the tables. Refer to Section 4.4.2 for details on the use of these parameters.

Table 24: Angular Orientations for ML Fit Rings

Geometry	Ring	$\psi_{\ell,1}$	$\psi_{\ell,2}$	$\psi_{\ell,3}$	$\psi_{\ell,4}$	$\psi_{\ell,5}$	$\psi_{\ell,6}$
2R	$\ell = 1$	0.50π	1.50π				
3R	$\ell = 1$	-0.17π	0.50π	1.17π			
4R	$\ell = 1$	0.00π	0.50π	1.00π	1.50π		
C-4R	$\ell = 1$	0.00π	0.50π	1.00π	1.50π		
C-5R	$\ell = 1$	0.00π	0.40π	0.80π	1.20π	1.60π	
C-6R	$\ell = 1$	0.00π	0.33π	0.67π	1.00π	1.33π	1.67π
C-5R-5R	$\ell = 1$	0.00π	0.40π	0.80π	1.20π	1.60π	
	$\ell = 2$	0.20π	0.60π	1.00π	1.40π	1.80π	
C-6R-6R	$\ell = 1$	0.00π	0.33π	0.67π	1.00π	1.33π	1.67π
	$\ell = 2$	0.17π	0.50π	0.83π	1.17π	1.50π	1.83π

Table 25: KL Divergence and Ring 1 Parameters for 2R

C_B	KL	ω_1	$\Delta\hat{r}_1$	$\Delta\hat{\theta}_1$	$\hat{\sigma}_{r,1}$	$\hat{\sigma}_{\hat{\theta},1}$	$\hat{\sigma}_{\psi,1}$	$\rho_{r\hat{\theta},1}$
0.10	0.001	0.333	-0.88	0.49	0.99	0.92	1.95	-0.19
0.20	0.007	0.333	-0.82	0.60	0.95	0.88	1.55	-0.26
0.30	0.018	0.333	-0.78	0.66	0.91	0.85	1.38	-0.29
0.40	0.035	0.333	-0.75	0.71	0.87	0.83	1.26	-0.31
0.50	0.058	0.333	-0.72	0.75	0.83	0.81	1.18	-0.32
0.60	0.087	0.333	-0.69	0.78	0.79	0.79	1.11	-0.33
0.70	0.120	0.333	-0.67	0.81	0.76	0.77	1.05	-0.32
0.80	0.156	0.333	-0.65	0.84	0.74	0.76	1.01	-0.32
0.90	0.195	0.333	-0.63	0.87	0.72	0.74	0.97	-0.32
1.00	0.234	0.333	-0.61	0.89	0.70	0.72	0.94	-0.32

Table 26: KL Divergence and Ring 1 Parameters for 3R

C_B	KL	ω_1	$\Delta\hat{r}_1$	$\Delta\hat{\theta}_1$	$\hat{\sigma}_{r,1}$	$\hat{\sigma}_{\hat{\theta},1}$	$\hat{\sigma}_{\psi,1}$	$\rho_{r\hat{\theta},1}$
0.10	0.001	0.333	−0.88	0.49	0.99	0.92	1.95	−0.19
0.20	0.007	0.333	−0.82	0.60	0.95	0.88	1.55	−0.26
0.30	0.018	0.333	−0.78	0.66	0.91	0.85	1.38	−0.29
0.40	0.035	0.333	−0.75	0.71	0.87	0.83	1.26	−0.31
0.50	0.058	0.333	−0.72	0.75	0.83	0.81	1.18	−0.32
0.60	0.087	0.333	−0.69	0.78	0.79	0.79	1.11	−0.33
0.70	0.120	0.333	−0.67	0.81	0.76	0.77	1.05	−0.32
0.80	0.156	0.333	−0.65	0.84	0.74	0.76	1.01	−0.32
0.90	0.195	0.333	−0.63	0.87	0.72	0.74	0.97	−0.32
1.00	0.234	0.333	−0.61	0.89	0.70	0.72	0.94	−0.32

Table 27: KL Divergence and Ring 1 Parameters for 4R

C_B	KL	ω_1	$\Delta\hat{r}_1$	$\Delta\hat{\theta}_1$	$\hat{\sigma}_{r,1}$	$\hat{\sigma}_{\hat{\theta},1}$	$\hat{\sigma}_{\psi,1}$	$\rho_{r\hat{\theta},1}$
0.30	0.010	0.250	−0.74	0.72	0.90	0.86	1.19	−0.29
0.40	0.020	0.250	−0.70	0.77	0.85	0.84	1.09	−0.32
0.50	0.035	0.250	−0.67	0.81	0.80	0.82	1.01	−0.34
0.60	0.054	0.250	−0.64	0.85	0.76	0.80	0.95	−0.35
0.70	0.077	0.250	−0.61	0.89	0.72	0.78	0.89	−0.35
0.80	0.103	0.250	−0.59	0.91	0.69	0.76	0.85	−0.36
0.90	0.131	0.250	−0.56	0.94	0.66	0.74	0.81	−0.36
1.00	0.162	0.250	−0.54	0.96	0.64	0.73	0.78	−0.36
1.10	0.194	0.250	−0.53	0.99	0.62	0.71	0.75	−0.36
1.20	0.226	0.250	−0.51	1.00	0.61	0.70	0.73	−0.36
1.30	0.259	0.250	−0.50	1.02	0.59	0.69	0.71	−0.37
1.40	0.293	0.250	−0.49	1.03	0.58	0.68	0.70	−0.37

Table 28: KL Divergence and Center Parameters for C-4R

C_B	KL	ω_c	$\Delta\hat{r}_c$	$\hat{\sigma}_{r,c}$	$\hat{\sigma}_{\alpha\varepsilon,c}$
0.40	0.013	0.369	−0.41	0.82	0.73
0.50	0.022	0.378	−0.40	0.75	0.72
0.60	0.034	0.382	−0.38	0.69	0.70
0.70	0.049	0.385	−0.37	0.63	0.68
0.80	0.066	0.385	−0.36	0.58	0.67
0.90	0.085	0.386	−0.35	0.54	0.65
1.00	0.105	0.387	−0.34	0.51	0.64
1.10	0.126	0.386	−0.33	0.48	0.62
1.20	0.149	0.385	−0.32	0.45	0.61
1.30	0.172	0.384	−0.31	0.43	0.60
1.40	0.195	0.383	−0.31	0.41	0.59
1.50	0.219	0.381	−0.30	0.39	0.58

Table 29: Ring 1 Parameters for C-4R

C_B	ω_1	$\Delta\hat{r}_1$	$\Delta\hat{\theta}_1$	$\hat{\sigma}_{r,1}$	$\hat{\sigma}_{\hat{\theta},1}$	$\hat{\sigma}_{\psi,1}$	$\rho_{r\hat{\theta},1}$
0.40	0.158	−0.75	1.09	0.85	0.77	0.80	−0.30
0.50	0.156	−0.72	1.14	0.81	0.75	0.75	−0.31
0.60	0.154	−0.68	1.18	0.77	0.73	0.71	−0.31
0.70	0.154	−0.65	1.22	0.73	0.71	0.67	−0.31
0.80	0.154	−0.62	1.25	0.70	0.69	0.65	−0.30
0.90	0.153	−0.59	1.28	0.67	0.67	0.62	−0.30
1.00	0.153	−0.57	1.30	0.65	0.65	0.61	−0.30
1.10	0.153	−0.55	1.32	0.62	0.64	0.59	−0.30
1.20	0.154	−0.53	1.34	0.60	0.63	0.58	−0.30
1.30	0.154	−0.51	1.35	0.59	0.61	0.57	−0.31
1.40	0.154	−0.50	1.36	0.57	0.60	0.56	−0.31
1.50	0.155	−0.49	1.37	0.56	0.60	0.55	−0.32

Table 30: KL Divergence and Center Parameters for C-5R

C_B	KL	ω_c	$\Delta\hat{r}_c$	$\hat{\sigma}_{r,c}$	$\hat{\sigma}_{\alpha\varepsilon,c}$
0.60	0.021	0.380	−0.45	0.70	0.69
0.70	0.031	0.378	−0.42	0.64	0.68
0.80	0.044	0.375	−0.40	0.59	0.66
0.90	0.058	0.373	−0.38	0.55	0.64
1.00	0.074	0.370	−0.36	0.51	0.63
1.10	0.091	0.368	−0.35	0.48	0.61
1.20	0.109	0.367	−0.34	0.45	0.60
1.30	0.128	0.365	−0.32	0.43	0.59
1.40	0.147	0.364	−0.31	0.41	0.58
1.50	0.167	0.363	−0.30	0.39	0.57
1.60	0.187	0.362	−0.30	0.37	0.56
1.70	0.208	0.362	−0.29	0.36	0.55
1.80	0.228	0.361	−0.28	0.34	0.55
1.90	0.249	0.360	−0.27	0.33	0.54

Table 31: Ring 1 Parameters for C-5R

C_B	ω_1	$\Delta\hat{r}_1$	$\Delta\hat{\theta}_1$	$\hat{\sigma}_{r,1}$	$\hat{\sigma}_{\hat{\theta},1}$	$\hat{\sigma}_{\psi,1}$	$\rho_{r\hat{\theta},1}$
0.60	0.124	−0.58	1.23	0.75	0.73	0.63	−0.34
0.70	0.124	−0.56	1.26	0.71	0.71	0.60	−0.34
0.80	0.125	−0.53	1.29	0.67	0.69	0.57	−0.34
0.90	0.125	−0.51	1.31	0.64	0.67	0.55	−0.34
1.00	0.126	−0.49	1.33	0.61	0.66	0.53	−0.34
1.10	0.126	−0.47	1.35	0.58	0.64	0.52	−0.34
1.20	0.127	−0.45	1.37	0.56	0.63	0.51	−0.35
1.30	0.127	−0.44	1.38	0.54	0.62	0.49	−0.35
1.40	0.127	−0.43	1.39	0.53	0.61	0.48	−0.35
1.50	0.127	−0.41	1.40	0.51	0.60	0.47	−0.36
1.60	0.128	−0.40	1.41	0.50	0.60	0.47	−0.36
1.70	0.128	−0.40	1.42	0.49	0.59	0.46	−0.37
1.80	0.128	−0.39	1.43	0.48	0.58	0.45	−0.37
1.90	0.128	−0.38	1.43	0.47	0.58	0.45	−0.38

Table 32: KL Divergence and Center Parameters for C-6R

C_B	KL	ω_c	$\Delta\hat{r}_c$	$\hat{\sigma}_{r,c}$	$\hat{\sigma}_{\alpha\varepsilon,c}$
0.80	0.034	0.377	−0.42	0.60	0.66
0.90	0.045	0.373	−0.40	0.55	0.64
1.00	0.058	0.369	−0.38	0.52	0.62
1.10	0.072	0.366	−0.36	0.48	0.61
1.20	0.088	0.363	−0.35	0.46	0.60
1.30	0.104	0.360	−0.33	0.43	0.58
1.40	0.121	0.358	−0.32	0.41	0.57
1.50	0.138	0.356	−0.31	0.39	0.56
1.60	0.156	0.355	−0.30	0.37	0.55
1.70	0.175	0.353	−0.29	0.36	0.55
1.80	0.193	0.352	−0.28	0.35	0.54
1.90	0.212	0.352	−0.28	0.33	0.53
2.00	0.231	0.352	−0.27	0.32	0.53
2.10	0.249	0.351	−0.26	0.31	0.52
2.20	0.268	0.351	−0.26	0.30	0.52

Table 33: Ring 1 Parameters for C-6R

C_B	ω_1	$\Delta\hat{r}_1$	$\Delta\hat{\theta}_1$	$\hat{\sigma}_{r,1}$	$\hat{\sigma}_{\hat{\theta},1}$	$\hat{\sigma}_{\psi,1}$	$\rho_{r\hat{\theta},1}$
0.80	0.104	−0.49	1.31	0.65	0.69	0.53	−0.35
0.90	0.104	−0.47	1.34	0.62	0.67	0.51	−0.36
1.00	0.105	−0.45	1.35	0.59	0.66	0.49	−0.36
1.10	0.106	−0.43	1.37	0.56	0.65	0.48	−0.36
1.20	0.106	−0.41	1.39	0.54	0.64	0.46	−0.37
1.30	0.107	−0.40	1.40	0.52	0.63	0.45	−0.37
1.40	0.107	−0.39	1.41	0.50	0.62	0.44	−0.37
1.50	0.107	−0.38	1.42	0.49	0.61	0.43	−0.38
1.60	0.108	−0.37	1.43	0.47	0.60	0.42	−0.38
1.70	0.108	−0.36	1.43	0.46	0.59	0.42	−0.39
1.80	0.108	−0.35	1.44	0.45	0.59	0.41	−0.39
1.90	0.108	−0.34	1.45	0.44	0.58	0.40	−0.40
2.00	0.108	−0.34	1.46	0.43	0.58	0.40	−0.40
2.10	0.108	−0.33	1.46	0.43	0.57	0.39	−0.41
2.20	0.108	−0.33	1.47	0.42	0.57	0.39	−0.41

Table 34: KL Divergence and Center Parameters for C-5R-5R

C_B	KL	ω_c	$\Delta\hat{r}_c$	$\hat{\sigma}_{r,c}$	$\hat{\sigma}_{\alpha\varepsilon,c}$
1.00	0.024	0.226	−0.28	0.54	0.55
1.20	0.037	0.219	−0.26	0.47	0.52
1.40	0.052	0.214	−0.23	0.41	0.50
1.60	0.069	0.207	−0.22	0.37	0.49
1.80	0.087	0.201	−0.20	0.33	0.47
2.00	0.106	0.196	−0.19	0.30	0.46
2.20	0.125	0.192	−0.18	0.27	0.44
2.40	0.145	0.188	−0.18	0.25	0.43
2.60	0.166	0.184	−0.17	0.24	0.42
2.80	0.187	0.181	−0.16	0.22	0.41
3.00	0.207	0.178	−0.15	0.21	0.41
3.20	0.228	0.175	−0.15	0.20	0.40
3.40	0.249	0.173	−0.14	0.19	0.39

Table 35: Ring 1 Parameters for C-5R-5R

C_B	ω_1	$\Delta\hat{r}_1$	$\Delta\hat{\theta}_1$	$\hat{\sigma}_{r,1}$	$\hat{\sigma}_{\hat{\theta},1}$	$\hat{\sigma}_{\psi,1}$	$\rho_{r\hat{\theta},1}$
1.00	0.130	−0.32	1.09	0.53	0.56	0.55	−0.18
1.20	0.129	−0.30	1.09	0.46	0.54	0.53	−0.17
1.40	0.129	−0.28	1.10	0.41	0.52	0.51	−0.16
1.60	0.128	−0.26	1.09	0.37	0.50	0.49	−0.15
1.80	0.128	−0.24	1.09	0.34	0.48	0.48	−0.15
2.00	0.127	−0.23	1.09	0.31	0.47	0.46	−0.14
2.20	0.127	−0.22	1.09	0.29	0.46	0.45	−0.14
2.40	0.126	−0.21	1.08	0.27	0.45	0.44	−0.13
2.60	0.126	−0.20	1.08	0.25	0.44	0.43	−0.13
2.80	0.126	−0.19	1.08	0.24	0.43	0.43	−0.13
3.00	0.125	−0.19	1.08	0.23	0.42	0.42	−0.12
3.20	0.125	−0.18	1.07	0.21	0.42	0.41	−0.12
3.40	0.124	−0.18	1.07	0.21	0.41	0.41	−0.12

Table 36: Ring 2 Parameters for C-5R-5R

C_B	ω_2	$\Delta\hat{r}_2$	$\Delta\hat{\theta}_2$	$\hat{\sigma}_{r,2}$	$\hat{\sigma}_{\hat{\theta},2}$	$\hat{\sigma}_{\psi,2}$	$\rho_{r\hat{\theta},2}$
1.00	0.024	−0.61	2.08	0.62	0.56	0.39	−0.23
1.20	0.027	−0.55	2.08	0.57	0.54	0.38	−0.24
1.40	0.028	−0.51	2.09	0.54	0.52	0.37	−0.25
1.60	0.030	−0.48	2.09	0.51	0.51	0.37	−0.25
1.80	0.032	−0.46	2.09	0.49	0.49	0.36	−0.27
2.00	0.033	−0.44	2.08	0.48	0.48	0.36	−0.28
2.20	0.035	−0.43	2.08	0.46	0.47	0.36	−0.29
2.40	0.036	−0.42	2.07	0.45	0.47	0.36	−0.30
2.60	0.037	−0.41	2.07	0.44	0.46	0.36	−0.31
2.80	0.038	−0.40	2.06	0.43	0.46	0.36	−0.32
3.00	0.039	−0.39	2.06	0.42	0.45	0.36	−0.33
3.20	0.040	−0.39	2.05	0.42	0.45	0.35	−0.33
3.40	0.041	−0.38	2.05	0.41	0.44	0.35	−0.34

Table 37: KL Divergence and Center Parameters for C-6R-6R

C_B	KL	ω_c	$\Delta\hat{r}_c$	$\hat{\sigma}_{r,c}$	$\hat{\sigma}_{\alpha\varepsilon,c}$
1.50	0.041	0.212	−0.23	0.39	0.49
1.70	0.054	0.206	−0.22	0.35	0.47
1.90	0.068	0.200	−0.20	0.32	0.46
2.10	0.082	0.195	−0.19	0.29	0.45
2.30	0.098	0.191	−0.18	0.27	0.43
2.50	0.114	0.187	−0.17	0.25	0.42
2.70	0.130	0.183	−0.17	0.23	0.41
2.90	0.147	0.179	−0.16	0.22	0.40
3.10	0.164	0.177	−0.15	0.21	0.40
3.30	0.181	0.174	−0.15	0.19	0.39
3.50	0.199	0.172	−0.14	0.18	0.38
3.70	0.216	0.169	−0.14	0.18	0.37
3.90	0.233	0.167	−0.13	0.17	0.37
4.10	0.250	0.166	−0.13	0.16	0.36

Table 38: Ring 1 Parameters for C-6R-6R

C_B	ω_1	$\Delta\hat{r}_1$	$\Delta\hat{\theta}_1$	$\hat{\sigma}_{r,1}$	$\hat{\sigma}_{\hat{\theta},1}$	$\hat{\sigma}_{\psi,1}$	$\rho_{r\hat{\theta},1}$
1.50	0.108	−0.25	1.12	0.39	0.50	0.46	−0.16
1.70	0.107	−0.23	1.12	0.35	0.49	0.44	−0.15
1.90	0.107	−0.22	1.11	0.32	0.47	0.43	−0.15
2.10	0.106	−0.21	1.11	0.29	0.46	0.42	−0.14
2.30	0.106	−0.20	1.11	0.27	0.45	0.41	−0.14
2.50	0.106	−0.19	1.10	0.25	0.44	0.40	−0.14
2.70	0.106	−0.18	1.10	0.24	0.43	0.40	−0.14
2.90	0.105	−0.17	1.10	0.22	0.42	0.39	−0.13
3.10	0.105	−0.17	1.09	0.21	0.42	0.38	−0.13
3.30	0.105	−0.16	1.09	0.20	0.41	0.38	−0.13
3.50	0.104	−0.16	1.09	0.19	0.40	0.37	−0.13
3.70	0.104	−0.15	1.08	0.18	0.40	0.37	−0.13
3.90	0.104	−0.15	1.08	0.18	0.39	0.36	−0.12
4.10	0.104	−0.15	1.08	0.17	0.39	0.36	−0.13

Table 39: Ring 2 Parameters for C-6R-6R

C_B	ω_2	$\Delta\hat{r}_2$	$\Delta\hat{\theta}_2$	$\hat{\sigma}_{r,2}$	$\hat{\sigma}_{\hat{\theta},2}$	$\hat{\sigma}_{\psi,2}$	$\rho_{r\hat{\theta},2}$
1.50	0.024	−0.43	2.14	0.49	0.51	0.32	−0.28
1.70	0.025	−0.40	2.14	0.46	0.50	0.32	−0.29
1.90	0.026	−0.38	2.13	0.44	0.49	0.31	−0.30
2.10	0.028	−0.37	2.13	0.42	0.48	0.31	−0.31
2.30	0.029	−0.35	2.13	0.41	0.47	0.31	−0.33
2.50	0.030	−0.34	2.12	0.39	0.46	0.31	−0.34
2.70	0.031	−0.34	2.12	0.38	0.46	0.31	−0.35
2.90	0.031	−0.33	2.11	0.37	0.45	0.30	−0.36
3.10	0.032	−0.32	2.11	0.37	0.45	0.30	−0.37
3.30	0.033	−0.32	2.10	0.36	0.44	0.30	−0.38
3.50	0.034	−0.31	2.10	0.35	0.44	0.30	−0.39
3.70	0.034	−0.31	2.09	0.35	0.44	0.30	−0.40
3.90	0.035	−0.31	2.09	0.34	0.44	0.30	−0.41
4.10	0.035	−0.30	2.08	0.34	0.43	0.30	−0.42

REFERENCES

[1] Alspach, D. L. and Sorenson, H. W., "Nonlinear Bayesian estimation using Gaussian sum approximations," *IEEE Transactions on Automatic Control*, vol. 17, no. 4, pp. 439–448, 1972.

[2] Bar-Shalom, Y., Li, X. R., and Kirubarajan, T., *Estimation with Applications to Tracking and Navigation: Theory Algorithms and Software*. John Wiley & Sons, 2001.

[3] Bilmes, J. A. and others, "A gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models," *International Computer Science Institute*, vol. 4, no. 510, p. 126, 1998.

[4] Coogle, R. A., Smith, L. D., and Blair, W. D., "Debiased coordinate conversion of bistatic radar measurements," in *Proceedings of the 2012 IEEE Radar Conference*, pp. 0383–0388, 2012.

[5] Coogle, R. A., Smith, L. D., and Blair, W. D., "Mitigating the bias in converted bistatic radar measurements using the unscented transform," in *Proceedings of the 2013 IEEE Aerospace Conference*, 2013.

[6] Crouse, D., "Basic tracking using nonlinear 3d monostatic and bistatic measurements," *IEEE Aerospace and Electronic Systems Magazine*, vol. 29, no. 8, pp. 4–53, 2014.

[7] Daum, F. and Huang, J., "Particle flow for nonlinear filters, Bayesian decisions and transport," in *Proceedings of the 16th International Conference on Information Fusion*, pp. 1072–1079, July 2013.

[8] Davis, B. and Blair, W. D., "Gaussian mixture approach to long range radar tracking with high range resolution," in *Proceedings of the 2015 IEEE Aerospace Conference*, 2015.

[9] Davis, B. and Blair, W. D., "Adaptive gaussian mixture modeling for tracking of long range targets," in *Proceedings of the 2016 IEEE Aerospace Conference*, 2016.

[10] Davis, B. and Blair, W. D., "Gaussian mixture modeling for bistatic measurements," in *Proceedings of the 2017 IEEE Aerospace Conference*, 2017.

[11] Davis, B. and Blair, W. D., "Three-dimensional tracking for long range radars using the measurement-adaptive gaussian mixture filter," *IEEE Transactions on Aerospace and Electronic Systems*, submitted Nov 2017.

[12] Davis, B. and Blair, W. D., "Long range radar targets tracking with optimal, gaussian mixture measurement models," *IEEE Transactions on Aerospace and Electronic Systems*, submitted Oct 2017.

[13] Duda, R. O., Hart, P. E., and Stork, D. G., *Pattern Classification (2nd Edition)*. Wiley-Interscience, 2000.

[14] Havlak, F. and Campbell, M., "Discrete and continuous, probabilistic anticipation for autonomous robots in urban environments," *IEEE Transactions on Robotics*, vol. 30, no. 2, pp. 461–474, 2014.

[15] Horwood, J. T. and Poore, A. B., "Adaptive gaussian sum filters for space surveillance," *IEEE Transactions on Automatic Control*, vol. 56, no. 8, pp. 1777–1790, 2011.

[16] Julier, S. J., Uhlmann, J. K., and Durrant-Whyte, H. F., "A new approach for filtering nonlinear systems," in *Proceedings of the 1995 American Control Conference*, vol. 3, pp. 1628–1632, 1995.

[17] Kalman, R. E., "A new approach to linear filtering and prediction problems," *Journal of basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[18] Kullback, S., *Information Theory and Statistics*. Dover Books on Mathematics, Dover Publications, 2012.

[19] Lerro, D. and Bar-Shalom, Y., "Tracking with debiased consistent converted measurements versus EKF," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 29, no. 3, pp. 1015–1022, 1993.

[20] Longbin, M., Xiaoquan, S., Yiyu, Z., Kang, S. Z., and Bar-Shalom, Y., "Unbiased converted measurements for tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 34, no. 3, pp. 1023–1027, 1998.

[21] Musicki, D. and Evans, R. J., "Measurement gaussian sum mixture target tracking," in *Proceedings of the 9th International Conference on Information Fusion*, 2006.

[22] Papoulis, A. and Pillai, S., *Probability, random variables, and stochastic processes*. McGraw-Hill electrical and electronic engineering series, McGraw-Hill, 2002.

[23] Romeo, K., Willett, P., and Bar-Shalom, Y., "Particle filter tracking for the banana problem," in *Proceedings of Signal and Data Processing of Small Targets 2013*, vol. 8857, SPIE, 2013.

[24] Romeo, K., Willett, P., and Bar-Shalom, Y., "Particle filter tracking for banana and contact lens problems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 51, no. 2, pp. 1098–1110, 2015.

[25] Salmond, D. J., "Mixture reduction algorithms for target tracking in clutter," in *Proceedings of Signal and Data Processing of Small Targets 1990*, vol. 1305, pp. 434–445, SPIE, 1990.

[26] Sorenson, H. W. and Alspach, D. L., "Recursive Bayesian estimation using Gaussian sums," *Automatica*, vol. 7, no. 4, pp. 465–479, 1971.

[27] Tian, X. and Bar-Shalom, Y., "Coordinate conversion and tracking for very long range radars," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, no. 3, pp. 1073–1088, 2009.

[28] Tian, X. and Bar-Shalom, Y., "A consistency-based gaussian mixture filtering approach for the contact lens problem," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, no. 3, pp. 1636–1646, 2014.

[29] Tian, X., Bar-Shalom, Y., Chen, G., Pham, K., and Blasch, E., "Track splitting technique for the contact lens problem," in *Proceedings of the 14th International Conference on Information Fusion*, 2011.

[30] Wan, E. A. and Van Der Merwe, R., "The Unscented Kalman filter," *Kalman Filtering and Neural Networks*, pp. 221–280, 2001.

[31] Willis, N., *Bistatic Radar*. Electromagnetics and Radar, Institution of Engineering and Technology, 2005.

[32] Zhang, Q. and Song, T. L., "Gaussian mixture measurements for very long range tracking," in *Proceedings of 12th International Conference on Informatics in Control, Automation and Robotics*, vol. 1, pp. 457–464, 2015.

[33] Zhang, Q. and Song, T. L., "Gaussian mixture presentation of measurements for long-range radar tracking," *Digital Signal Processing*, vol. 56, pp. 110–122, 2016.

VITA

Benjamin Davis was born in Melbourne, FL in 1986. He received his B.S. and M.S. degrees in Electrical Engineering from the Georgia Institute of Technology in 2010 and 2011, with a B.S. Minor in Computer Science. During his undergraduate studies he worked as a software engineer with Datapath, Inc., now part of Rockwell Collins. He is currently employed as a research engineer at the Georgia Tech Research Institute (GTRI) Sensors and Electromagnetics Applications Laboratory in Smyrna, GA. He has 7 years of experience supporting US Army tactical sensors and systems as an estimation and tracking specialist, including adaptation of tracking benchmark software to more effectively model state-of-the art tactical sensors. In his spare time, Ben enjoys hiking, biking, and playing the piano and guitar. He plans to receive his PhD in Electrical and Computer Engineering from Georgia Tech in Dec 2017.